

PATENT ABSTRACTS OF JAPAN

(11)Publication number : 2003-004739

(43)Date of publication of application : 08.01.2003

(51)Int.Cl.

G01N 33/53
 C12M 1/00
 C12N 15/09
 C12Q 1/68
 G06F 17/30
 G06N 3/12

(21)Application number : 2001-192146

(71)Applicant :

TAKARA BIO INC
 MITSUBISHI SPACE SOFTWARE KK

(22)Date of filing : 26.06.2001

(72)Inventor :

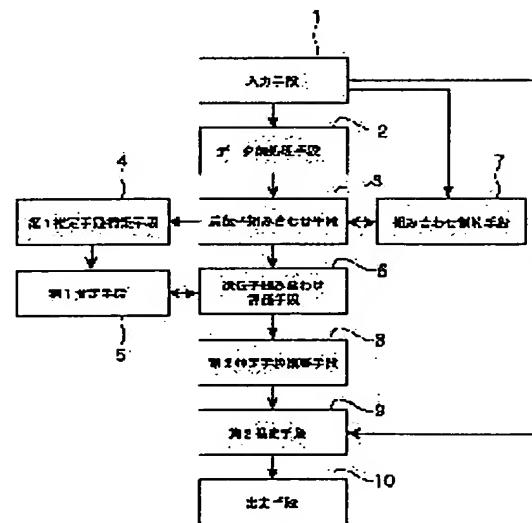
YOSHIKAWA YOSHIE
 MINENO JUNICHI
 MUKAI HIROYUKI
 ASADA KIYOSU
 KATO IKUNOSHIN
 KAGAWA YOSHINORI
 YAJIMA SHIGEKI
 NAGANO TAKAFUMI
 YAMADA SATOSHI

(54) METHOD AND APPARATUS FOR ANALYZING GENE MANIFESTATION PROFILE

(57)Abstract:

PROBLEM TO BE SOLVED: To estimate target evaluation indicator data from gene manifestation profile data.

SOLUTION: A gene combination candidate is generated, and an evaluation indicator is estimated with the use of the generated gene combination. The gene combination is evaluated on the basis of the estimation result. The generation of the gene combination is repeatedly carried out on the basis of the evaluation. The gene combination not lower than a set evaluation is obtained accordingly. A target evaluation indicator is estimated with the use of the gene combination.



LEGAL STATUS

[Date of request for examination]

[Date of sending the examiner's decision of rejection]

[Kind of final disposal of application other than the examiner's decision of rejection or application converted registration]

[Date of final disposal for application]

[Patent number]

[Date of registration]

[Number of appeal against examiner's decision of rejection]

[Date of requesting appeal against examiner's decision of rejection]

[Date of extinction of right]

This Page Blank (uspto)

**Japan se Laid-Open Pat nt Publication
No. 4739/2003 (Tokukai 2003-4739)**

A. Relevance of the Above-identified Document

The following is a partial English translation of exemplary portions of non-English language information that may be relevant to the issue of patentability of the claims of the present application.

B. Translation of the Relevant Passages of the Document

See also the attached English Abstract.

[PROBLEMS TO BE SOLVED BY THE INVENTION]

[0008]

Generally, genetic diseases involve complex interplay of different genes, and individual gene expression data do not always show a good correlation with the evaluation index. For these reasons, it has not been possible to provide an accurate estimate with the gene expression profile data that have been selected based on correlation coefficients. That is, it has not been possible to provide a good estimate unless genes are properly selected.

[0009]

The present invention was made to solve the foregoing problems, and it is an object of the present invention to provide a gene expression profile analyzing method and a gene expression profile analyzing apparatus, with which, for the estimation of an evaluation index, genes are selected that are suitable for quantitatively estimating an evaluation index of interest from data obtained from each subject.

[0010]

[MEANS TO SOLVE THE PROBLMES]

This Page Blank (uspto)

A gene expression profile analyzing apparatus according to the present invention includes: input means for inputting data; gene combining means for combining genes; first-estimation-means constructing means for constructing means for estimating evaluation index data; first estimation means for estimating evaluation index data; gene combination evaluating means for evaluating a combination of generated genes; and output means for outputting the combination of genes and an estimation method.

[0011]

The gene expression profile analyzing apparatus further includes data pre-processing means for removing unnecessary genes or normalizing data.

[0012]

The gene expression profile analyzing apparatus further includes: means for constructing second estimation means that estimates evaluation estimation data using a combination of optimum genes; and second estimation means.

[0013]

The gene expression profile analyzing apparatus further includes combination restricting means for restricting the combination made by the gene recombining means.

[0014]

The gene combining means employs genetic algorithm.

[0015]

The combination evaluation in the genetic algorithm employed by the gene recombining means is based on an estimated error of the evaluation index data and the number of employed genes.

[0016]

As a method of changing combinations of genes in the genetic algorithm employed by the gene recombining means,

This Page Blank (uspto)

genes are adopted in substantially the same number as the genes that are rejected.

[0017]

As the second estimation means constructing means, a neural network employing error back propagation learning is used.

[0018]

A gene expression profile analyzing method of the invention includes: the input step of inputting a gene expression profile and evaluation index data of interest by converting these data into input data; the gene combination candidate generating step, in which one or more gene combinations ("gene combination groups") made up of arbitrary numbers of genes are generated; the estimation method generating step, in which an estimation method of evaluation index is generated using the gene combination groups; the estimation step, in which an estimate value of evaluation index is determined using the estimation method and the gene combination groups; and the gene combination extracting step, in which the estimate value of estimation index is used and if there is one or more gene combination groups that are closely related to the estimation index, then the gene combination groups are extracted, and in which if the estimate value that has the best evaluation in a group of estimate values for the extracted gene combination groups is below a preset value, then the procedure from the gene combination candidate generating step to the estimation step is repeated, and if not, the extracted gene combination groups and the estimation method are outputted in the output step.

[0019]

In the gene expression profile analyzing method, the input gene expression profile is normalized, and a gene expression

This Page Blank (uspto)

profile that has a small correlation with the evaluation index data of interest is removed.

[0020]

The gene expression profile analyzing method further includes the second estimation method generating step, in which an estimation method is generated that, with the extracted gene combination groups and estimation method groups outputted in the output step, estimates evaluation index data corresponding to the gene expression profile and the estimation methods. In the second estimation method generating step, an estimation method is generated, and, with the estimation method so generated, an estimate value of evaluation index data for the gene expression profile is obtained.

[0021]

The gene expression profile analyzing method further includes the gene combination restricting step, in which a restriction is placed on combinations of genes in the evaluation of the gene combination groups, using known information concerning gene functions or relations between genes, so as to change the estimate value for the evaluation index data of the gene combination groups.

[0022]

The gene combination generating method employs genetic algorithm.

[0023]

The evaluation of gene combinations is based on an estimated error of evaluation index data, and the number of employed genes.

[0024]

As the method of varying gene combinations in the genetic algorithm, genes are adopted in substantially the same number

This Page Blank (uspto)

as the genes that are rejected.

[0025]

As the second estimation method generating step for generating an estimation method used to estimate evaluation index data, an estimation method is learned using a neural network employing error back propagation learning.

[0026]

In storing the gene combination groups and the estimation method groups outputted in the manner described above, database is used in which at least the gene combinations, the evaluation index, and the estimation method for the estimation index data are stored by being associated with one another.

[0027]

Further, database is used in which the evaluation index group and the estimation method groups for the evaluation index are stored by being associated with the gene combination groups.

This Page Blank (uspto)

(19)日本国特許庁 (JP)

(12) 公開特許公報 (A)

(11)特許出願公開番号

特開2003-4739

(P2003-4739A)

(43)公開日 平成15年1月8日(2003.1.8)

(51)Int.Cl. ⁷	識別記号	F I	テマコード(参考)
G 0 1 N 33/53		G 0 1 N 33/53	Z 4 B 0 2 4
C 1 2 M 1/00		C 1 2 M 1/00	A 4 B 0 2 9
C 1 2 N 15/09		C 1 2 Q 1/68	A 4 B 0 6 3
	Z N A	G 0 6 F 17/30	1 7 0 F 5 B 0 7 5
C 1 2 Q 1/68		G 0 6 N 3/12	

審査請求 未請求 請求項の数40 O L (全 76 頁) 最終頁に続く

(21)出願番号 特願2001-192146(P2001-192146)

(71)出願人 302019245

タカラバイオ株式会社

滋賀県大津市瀬田三丁目4番1号

(22)出願日 平成13年6月26日(2001.6.26)

(71)出願人 591102095

三菱スペース・ソフトウエア株式会社

東京都港区浜松町2丁目4番1号

(74)上記1名の代理人 100104776

弁理士 佐野 弘

(72)発明者 吉川 良恵

滋賀県大津市瀬田三丁目4番1号 寶酒造

株式会社中央研究所内

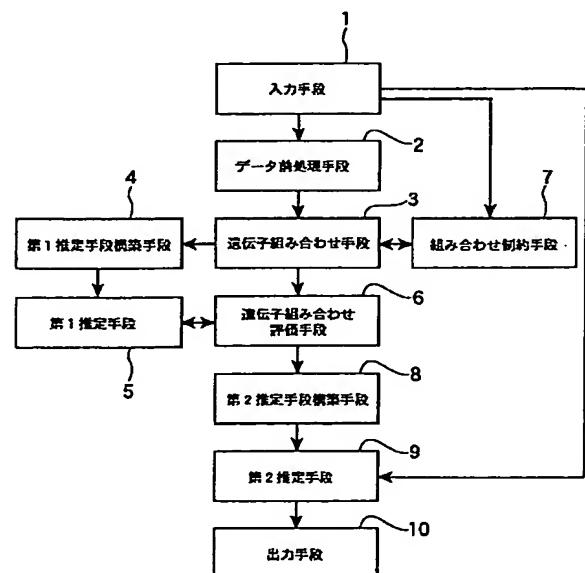
最終頁に続く

(54)【発明の名称】 遺伝子発現プロファイル解析方法及び装置

(57)【要約】

【課題】 遺伝子発現プロファイルデータから注目する評価指標データを推定する。

【解決手段】 遺伝子の組み合わせ候補を生成し、生成された遺伝子の組み合わせを用いて評価指標を推定し、その推定結果に基づき遺伝子の組み合わせを評価し、その評価に基づき遺伝子の組み合わせを生成することを繰り返すことによって、設定した評価以上の遺伝子の組み合わせを求める、その遺伝子の組み合わせを用いて注目する評価指標を推定する構成にした。



【特許請求の範囲】

【請求項1】各試料における遺伝子発現プロファイルデータと注目する評価指標データとを受け取る入力手段と、入力されたデータから評価指標データ推定のための遺伝子の組み合わせ候補を生成する遺伝子組み合わせ手段と、遺伝子の組み合わせ候補の遺伝子発現プロファイルデータから評価指標データを推定する手段を構築する第1推定手段構築手段と、該第1推定手段構築手段により構築され遺伝子の組み合わせ候補の遺伝子発現プロファイルデータから評価指標データを推定する第1推定手段と、前記第1推定手段を用いて前記遺伝子の組み合わせ候補を評価し、予め設定された基準を満たしていれば組み合わせ結果を出力手段に出力し、満たしていなければ再度前記遺伝子組み合わせ手段に処理を戻す遺伝子組み合わせ評価手段と、前記遺伝子の組み合わせ結果と評価指標データを推定する推定法を出力する前記出力手段とを備え、注目する評価指標と関係の深い遺伝子を抽出し、評価指標データを推定する推定法を得ることを特徴とする遺伝子発現プロファイル解析装置。

【請求項2】遺伝子発現プロファイルデータを正規化、及び注目する評価指標と相関の低い遺伝子発現データを取り除く処理ができるデータ前処理手段を備えたことを特徴とする請求項1に記載の遺伝子発現プロファイル解析装置。

【請求項3】遺伝子の組み合わせ結果の遺伝子発現プロファイルデータから評価指標データを推定する手段を構築する第2推定手段構築手段と、該第2推定手段構築手段により構築され前記遺伝子の組み合わせ結果の遺伝子発現プロファイルデータから評価指標データを推定する第2推定手段とを備え、前記遺伝子の組み合わせ結果として抽出された遺伝子発現プロファイルデータを受け取り、前記第2推定手段を利用して評価指標データを推定できるように構成したことを特徴とする請求項1又は2に記載の遺伝子発現プロファイル解析装置。

【請求項4】遺伝子の組み合わせ手段における組み合わせを制約する組み合わせ制約手段を備え、遺伝子の機能、遺伝子同士の関連に関する既知の知識を利用できることを特徴とする請求項1乃至3のいずれか1項に記載の遺伝子発現プロファイル解析装置。

【請求項5】遺伝子の組み合わせ手段に遺伝的アルゴリズムを用いることを特徴とする請求項1乃至4のいずれか1項に記載の遺伝子発現プロファイル解析装置。

【請求項6】遺伝子の組み合わせ手段に用いる遺伝的アルゴリズムで遺伝子の組み合わせを評価する指標として、評価指標データの推定誤差と採用する遺伝子数に基づく指標を用いることを特徴とする請求項5記載の遺伝子発現プロファイル解析装置。

【請求項7】遺伝子の組み合わせ手段に用いる遺伝的アルゴリズムにおける遺伝子の組み合わせの変化法として、遺伝子の採用と非採用を同程度変化させ、採用する

遺伝子数を大きく変化させないことを特徴とする請求項5又は6に記載の遺伝子発現プロファイル解析装置。

【請求項8】第2推定手段構築手段として、誤差逆伝播学習を用いたニューラルネットワークによって推定法を学習することを特徴とする請求項4又は7記載の遺伝子発現プロファイル解析装置。

【請求項9】(1) 遺伝子発現プロファイルデータと評価指標データを入力するステップ、(2) 入力されたデータから遺伝子の組み合わせ候補を生成し、評価指標データを推定する方法を生成するステップ、(3)

(2) で生成された方法で評価指標データを推定するステップ、及び(4) (3) で得られた遺伝子の組み合わせ候補の評価指標の推定値を用いて、評価指標と関係の深い一つ又は複数の遺伝子を抽出するステップを含む、注目する評価指標と関係の深い一つ又は複数の遺伝子の組み合わせと、評価指標データを推定する推定法を得ることを特徴とする遺伝子発現プロファイル解析方法。

【請求項10】(1) 遺伝子発現プロファイルデータを正規化するステップ、及び(2) 注目する評価指標と相関の低い遺伝子発現データを取り除くステップを含むデータ前処理ステップを含むことを特徴とする請求項9に記載の遺伝子発現プロファイル解析方法。

【請求項11】請求項9又は10に記載の方法を用いて抽出された、遺伝子の組み合わせと評価指標データと評価指標データを推定する推定法と遺伝子の組み合わせの遺伝子発現プロファイルデータを用いて、(1) 抽出された遺伝子組み合わせとその遺伝子発現プロファイルデータからその評価指標データを推定する推定法を生成するステップ、及び(2) (1) で生成された推定法で評価指標データを推定するステップを含み、評価指標データを推定できることを特徴とする遺伝子発現プロファイル解析方法。

【請求項12】(1) 遺伝子の機能、遺伝子同士の関連に関する既知の知識を入力するステップ、及び(2)

(1) で入力された既知の知識から遺伝子の組み合わせを制約するステップを含む、請求項9乃至11のいずれか1項に記載の遺伝子発現プロファイル解析方法。

【請求項13】遺伝子の組み合わせの生成方法に、遺伝的アルゴリズムを用いることを特徴とする請求項9乃至12のいずれか1項に記載の遺伝子発現プロファイル解析方法。

【請求項14】遺伝子の組み合わせを評価する指標として、前記評価指標データの推定誤差と採用する遺伝子の数に基づく指標を用いることを特徴とする請求項9乃至13のいずれか1項に記載の遺伝子発現プロファイル解析方法。

【請求項15】遺伝的アルゴリズムにおける遺伝子の組み合わせの変化法として、遺伝子の採用と非採用を同程度変化させ、採用する遺伝子数を大きく変化させないこ

とを特徴とする請求項9乃至14のいずれか1項に記載の遺伝子発現プロファイル解析方法。

【請求項16】評価指標データを推定する推定法を生成する方法として、誤差逆伝播学習を用いたニューラルネットによって推定法を学習することを特徴とする請求項12又は15に記載の遺伝子発現プロファイル解析方法。

【請求項17】試料数nの被検試料における遺伝子発現プロファイルデータと、当該試料の評価指標データから評価指標データの推定に使用できる遺伝子を選択し、当該選択された遺伝子の発現データから各被検試料の評価指標データを推定する方法において、評価指標データ推定のための遺伝子の組み合わせ候補の数をn未満とし、評価指標データを推定する推定手段により得られる評価指標データの推定誤差から遺伝子の組み合わせ候補を評価することを特徴とする評価指標データの推定方法。

【請求項18】請求項9乃至17のいずれか1項に記載の方法を用いて選択された遺伝子、提供される評価指標データの推定法に基づいて被検試料における評価指標データを推定する方法であって、(1)被検試料における、選択された遺伝子発現プロファイルデータを得る工程、及び(2)工程1で得られた遺伝子発現プロファイルデータを前記評価指標データの推定法で処理し、被検試料の評価指標データを推定する工程を包含することを特徴とする被検試料の評価指標データ推定方法。

【請求項19】工程(1)において、選択された遺伝子発現プロファイルデータが被検試料中の遺伝子由来のmRNA量の測定から得られることを特徴とする請求項18に記載の評価指標データ推定方法。

【請求項20】mRNA量の測定がハイブリダイゼーション法又は核酸増幅法により実施される請求項19に記載の評価指標データ推定方法。

【請求項21】mRNA量の測定がDNAアレイを使用したハイブリダイゼーション法により実施される請求項20に記載の評価指標データ推定方法。

【請求項22】被検試料が被験者由来の癌病変部試料であり、評価指標データが前記被験者の予後である請求項18乃至21のいずれか1項に記載の評価指標データ推定方法。

【請求項23】請求項9乃至17のいずれか1項に記載の方法を用いて選択された遺伝子、又はその断片が担体上の予め定められた領域に固定化されていることを特徴とするDNAアレイ。

【請求項24】請求項9乃至17のいずれか1項に記載の方法を用いて選択された遺伝子、又はその断片にそれぞれハイブリダイズ可能なプローブを含有することを特徴とするキット。

【請求項25】請求項9乃至17のいずれか1項に記載の方法を用いて選択された遺伝子、又はその断片を核酸増幅法によってそれぞれ増幅するための少なくとも1種

のプライマーを含有することを特徴とするキット。

【請求項26】被検試料における遺伝子発現プロファイルデータより当該試料に関する評価指標データを得るために装置であって、(1)請求項9乃至17のいずれか1項に記載の方法を用いて選択された前記遺伝子発現プロファイルデータを入力する手段、(2)請求項9乃至17のいずれか1項に記載の方法を用いて提供される前記評価指標データの推定法により、前記手段(1)で得られた前記遺伝子発現プロファイルデータから被検試料における前記評価指標データを推定する手段、及び(3)前記手段(2)で得られた評価指標データを出力する手段、を包含することを特徴とする評価指標データ推定装置。

【請求項27】手段(1)が、請求項9乃至17のいずれか1項に記載の方法を用いて選択された遺伝子の発現を測定する手段を備えていることを特徴とする請求項26に記載の評価指標データ推定装置。

【請求項28】手段(1)が、選択された遺伝子、又はその断片が担体上の予め定められた領域に固定化されたDNAアレイを備えていることを特徴とする請求項27に記載の評価指標データ推定装置。

【請求項29】遺伝子組み合わせの評価に数式

$$f = \alpha f_{error} + (1 - \alpha) f_{count}$$

を用いたことを特徴とする請求項1に記載の遺伝子プロファイル解析装置。

【請求項30】遺伝子の採用、非採用の決定に数式

$$\lambda_{out} |G_{in}| = \lambda_{in} |G_{out}|$$

を用いたことを特徴とする請求項1に記載の遺伝子プロファイル解析装置。

【請求項31】遺伝子組み合わせの評価に数式

$$f = \alpha f_{error} + (1 - \alpha) f_{count}$$

を用いたことを特徴とする請求項9に記載の遺伝子プロファイル解析方法。

【請求項32】遺伝子の採用、非採用の決定に数式

$$\lambda_{out} |G_{in}| = \lambda_{in} |G_{out}|$$

を用いたことを特徴とする請求項15に記載の遺伝子プロファイル解析方法。

【請求項33】請求項1乃至22又は請求項26乃至32のいずれか1項に記載の方法又は装置により抽出された評価指標と関係の深い遺伝子発現プロファイルデータと評価指標データの推定法のデータを蓄積し、それらの中の少なくとも1つ以上の遺伝子発現プロファイルデータ及び評価指標データの推定法のデータを用いて被検試料の遺伝子発現プロファイルデータを分類する方法。

【請求項34】請求項1乃至22又は請求項26乃至33のいずれか1項に記載の方法又は装置により抽出され

た評価指標と関係の深い遺伝子発現プロファイルデータと評価指標データの推定法のデータと遺伝子の組み合わせを蓄積し、それらの中の少なくとも1つ以上の遺伝子発現プロファイルデータ及び評価指標データの推定法データを用いて被検試料の遺伝子発現プロファイルデータを分類する方法。

【請求項35】請求項1乃至22又は請求項26乃至33のいずれか1項に記載の方法又は装置を用いて選択された遺伝子と提供された評価指標データの推定法を蓄積する際に、少なくとも遺伝子と評価指標データと評価指標データの推定法を関連付けて蓄積することを特徴とするデータベース。

【請求項36】遺伝子の組み合わせに対して、評価指標データ及び評価指標データの推定法を関連付けて蓄積することを特徴とする請求項35に記載のデータベース。

【請求項37】請求項35又は36に記載のデータベースを備えていることを特徴とする請求項1に記載の遺伝子プロファイル解析装置。

【請求項38】請求項35又は36に記載のデータベースを用いることを特徴とする請求項9に記載の遺伝子プロファイル解析方法。

【請求項39】評価指標が被検試料を得た被験者の予後である、請求項1に記載の遺伝子発現プロファイル解析装置。

【請求項40】評価指標が被検試料を得た被験者の予後である、請求項9に記載の遺伝子発現プロファイル解析方法。

【発明の詳細な説明】

【0001】

【発明の属する技術分野】この発明は、DNAチップ、DNAマイクロアレイで計測した遺伝子発現プロファイルデータを解析し、評価指標の推定に有用な解析方法及び解析装置等に関する。

【0002】

【従来の技術】従来、病気の分類は被検試料の諸症状や病変部位の形態学的な差異によってなされていた。例えば、胃癌の場合には、癌の深達度・リンパ節転移の有無・他の臓器への転移の有無などの形態学的な所見によって胃癌の病期（ステージ）を判断していた。しかし、形態学的な違いによる分類には明確な判断基準がないことにより、しばしば不正確であり、客観的な分類が困難な場合が多い。また複数の病理診断医間で判定が異なる場合も多い。

【0003】全ての細胞はその生物の一そろいの遺伝子を持っているが、細胞の種類・時期によって発現されている遺伝子の種類・量は異なる。各細胞・組織で発現されている遺伝子の種類・量のパターンを遺伝子発現プロファイルと呼んでいる。各細胞の機能・性質はその時点での細胞内にある蛋白質の種類と分布によって決定される。従って、蛋白質の合成量を計測できる遺伝子発現プロ

ファイルの計測から、その細胞の機能・性質を推定することが可能であると考えられる。

【0004】疾病（特に癌）によって、遺伝子発現プロファイルは正常な細胞の遺伝子発現プロファイルから大きく変化することが知られている。逆に言えば、遺伝子発現プロファイルを計測することによって、病気の状態を遺伝子発現のレベルの面から客観的に分類することが可能になると考えられる。また、病気に伴う性質（癌の転移しやすさ、手術後の予後等）も遺伝子発現プロファイルの変化から定量的に推定することが可能になると考えられる。近年、多数の遺伝子をチップ上に固定し、同時に多数の遺伝子の発現を計測できるDNAチップ・DNAマイクロアレイが開発されてきており、遺伝子発現プロファイルを計測することが可能になってきた。

【0005】DNAチップで計測された多数の遺伝子発現量の変化のデータ解析法として、自己組織化写像を用いてデータをクラスター化して分類する方法（特開2000-342299号）や特定の疾病に関するクラス分けに有効な遺伝子を相関解析から選択し、重み付け票決スキームでクラス分けする方法（特開2001-017171号）がある。また、サポートベクトルマシン（Support Vector Machine: SVM）を用いて遺伝子の分類を学習する方法がある（Proceedings of National Academy of Science USA (2000) 97、第262頁～第267頁）。

【0006】疾病に伴う遺伝子発現プロファイルの変化を計測する場合、被検試料を多量に集めることが困難があるので、発現量の変化を計測する遺伝子の数と比較して、被検試料の数が非常に少ないという問題がある。被検試料の数が非常に少ないので、通常の統計的な手法では疾病との相関解析が困難である場合が多い。

【0007】

【発明が解決しようとする課題】しかしながら、以上のような従来の技術では、各被検試料をクラス分けする方法はあるが、各被検試料から得られるデータから注目する評価指標（例えば、手術後の予後）を定量的に予測できないという問題点があった。定量的に予測する場合、計測した全ての遺伝子の発現プロファイルデータを用いて予測すると、不必要的データも含まれているため推定精度が悪い、遺伝子の数に対して被検試料の数が圧倒的に少ないと推定に汎用性がないといった問題があった。

【0008】一般に遺伝子と関係する疾病においては複数の遺伝子が複雑に関係していることが多く、かつ個々の遺伝子発現データ自体と評価指標とは必ずしも高い相関を示すわけではないため、相関係数に基づいて選択した、遺伝子の発現プロファイルデータを用いた推定では精度のよい推定ができなかった。したがって、適切な遺伝子を選択しないとよい推定ができないという問題があ

った。

【0009】この発明は、以上のような課題を解決するためになされたもので、各被検試料から得られるデータから注目する評価指標を定量的に推定するための適切な遺伝子を選択し、評価指標を推定する遺伝子発現プロファイル解析方法及び解析装置等を提供することを目的とする。

【0010】

【課題を解決するための手段】この発明に係わる遺伝子発現プロファイル解析装置においては、データを入力する入力手段と、遺伝子の組み合わせを生成する遺伝子組み合わせ手段と、評価指標データを推定するための手段を構築する第1推定手段構築手段と、評価指標データを推定する第1推定手段と、生成した遺伝子の組み合わせを評価する遺伝子組み合わせ評価手段と、前記遺伝子の組み合わせと推定法を出力する出力手段とを備えたものである。

【0011】また、不必要的遺伝子を除いたり、データの正規化などを行うデータ前処理手段を備えたものである。

【0012】また、求めた最適の遺伝子の組み合わせを用いて評価指標データを推定する第2推定手段構築手段と第2推定手段を備えたものである。

【0013】また、前記遺伝子組み合わせ手段における組み合わせを制約する組み合わせ制約手段を備えたものである。

【0014】また、前記遺伝子の組み合わせ手段に遺伝的アルゴリズムを用いたものである。

【0015】また、前記遺伝子の組み合わせ手段に用いた前記遺伝的アルゴリズムの組み合わせの評価に評価指標データの推定誤差と採用する遺伝子数に基づく評価を用いるものである。

【0016】また、前記遺伝子の組み合わせ手段に用いる前記遺伝的アルゴリズムの遺伝子の組み合わせの変化法として、遺伝子の採用と非採用を同程度変化させるようにしたものである。

【0017】また、前記第2推定手段構築手段として、誤差逆伝播学習を用いたニューラルネットを用いたものである。

【0018】この発明における遺伝子発現プロファイル解析方法は、遺伝子発現プロファイルと注目する評価指標データを入力する入力ステップにて、入力データに変換し、次に、遺伝子組み合わせ候補生成ステップにて、任意の数の遺伝子組み合わせで構成される1つ以上の遺伝子の組み合わせ（以下、遺伝子組み合わせ群と称す）を生成し、次に、推定方法生成ステップにて生成した遺伝子組み合わせ群を用いて評価指標を推定する方法を生成し、推定ステップにて生成した推定方法と遺伝子組み合わせ群を用いて評価指標の推定値を求め、評価指標の推定値を用いて遺伝子組み合わせ抽出ステップにて評価

指標と関係の深い遺伝子組み合わせ群が1つ以上有ればそれらの遺伝子組み合わせ群を抽出し、抽出した遺伝子組み合わせ群の推定値群の中で最もよい評価があらかじめ設定した値よりよくない場合は遺伝子組み合わせ候補ステップから推定ステップまでを繰り返し、よい場合は、出力ステップにて抽出された遺伝子組み合わせ群と推定法を出力することを特徴としている。

【0019】また、入力した遺伝子発現プロファイルを正規化し、注目する評価指標データと相關の低い遺伝子発現プロファイルを取り除くことを特徴としている。

【0020】また、前記出力ステップにて抽出された遺伝子組み合わせ群と推定法群を用いて遺伝子発現プロファイルと推定法に対応した評価指標データを推定する方法を生成する第2推定方法生成ステップで推定方法を生成し、生成した推定方法を用いて遺伝子発現プロファイルの評価指標データの推定値を得ることを特徴としている。

【0021】また、前記遺伝子組み合わせ群の評価において遺伝子の機能、遺伝子同士の関連などに関する既知の知識を用いて、遺伝子の組み合わせに制約を加えて、遺伝子組み合わせ群の評価指標データ推定値を変化させる遺伝子組み合わせ制約ステップを特徴としている。

【0022】また、前記遺伝子組み合わせ生成方法に遺伝的アルゴリズムを用いることを特徴としている。

【0023】また、前記遺伝子組み合わせの評価に評価指標データの推定誤差と採用する遺伝子数に基づく評価を用いることを特徴としている。

【0024】また、前記遺伝的アルゴリズムにおける遺伝子の組み合わせの変化法として、遺伝子の採用と非採用を同程度変化させることを特徴としている。

【0025】また、前記評価指標データを推定する推定法を生成する第2推定方法生成ステップとして、誤差逆伝播学習を用いたニューラルネットによって推定法を学習することを特徴としている。

【0026】また、上記の方法にて出力した遺伝子組み合わせ群と推定法群を蓄積する際に、少なくとも遺伝子組み合わせと評価指標と評価指標データの推定法を関連付けて蓄積するデータベースを用いることを特徴としている。

【0027】また、前記遺伝子組み合わせ群に対して、評価指標群及び評価指標の推定法群を関連付けて蓄積するデータベースを用いることを特徴としている。

【0028】

【発明の実施の形態】【発明の実施の形態1】この明細書において、「遺伝子発現プロファイルデータ」とは、個々の試料、例えば組織、細胞等において発現されている複数の遺伝子の発現パターンを指し、言い換えれば遺伝子の種類とそのそれぞれの発現量（若しくは発現比率）から構成されるデータの集合体を意味する。母集団から選択された遺伝子群に関する発現パターンもこの発

明に言う遺伝子発現プロファイルデータに包含される。また、「遺伝子発現データ」とは、個々の試料における任意の遺伝子の発現量（若しくは発現比率）を指す。

【0029】この明細書に記載の「評価指標」とは、試料の性格付けに関連する任意の性質を指し、定性的な指標、定量的な指標のいずれもが包含される。この発明に使用される評価指標には特に限定はないが、例えば、疾患に関連するものでは疾患の名称、原因、進行状況、予後、余命や発症、再発、転移の可能性等が挙げられる。また、「評価指標データ」とは、各試料における特定の評価指標に関するデータを指す。すなわち、各試料が特定の評価指標に関してどのような数値を有するか、或いはどのようなクラスに分類されるかを示すデータである。

【0030】この発明において使用される試料としては、評価指標データが既知のもの、未知のもののいずれもが包含される。例えば、既知の評価指標データをこの発明に使用し、当該評価指標データの推定に有用な遺伝子群、並びに該遺伝子群の発現プロファイルデータから評価指標データを推定する方法を見出すことができる。

また、こうして得られた推定法と評価指標データ未知の試料における遺伝子発現プロファイルデータから、当該試料に関する評価指標データを推定することができる。

【0031】図1は、この発明の実施の形態1の遺伝子発現プロファイル解析装置の構成を示すブロック図である。

【0032】図1において、1は遺伝子発現プロファイルデータと注目する評価指標データを入力する入力手段、2は遺伝子発現プロファイルデータの正規化、注目する評価指標データとの相関の低い遺伝子データの削除などを行うデータ前処理手段、3は推定に用いる遺伝子の組み合わせを選択する遺伝子組み合わせ手段、4は選択された遺伝子の発現データから評価指標データを推定する手段を構築する推定手段構築手段、5は第1推定手段構築手段4で構築された方法により評価指標データを推定する第1推定手段、6は得られた推定法があらかじめ設定した基準を満たしているかどうかを評価する遺伝子組み合わせ評価手段、7は遺伝子の機能・関連などの既知の情報により遺伝子の組み合わせに制約を設ける組み合わせ制約手段、8は得られた遺伝子の組み合わせから新たに評価指標データを推定する手段を構築する第2推定手段構築手段、9は第2推定手段構築手段8で構築された推定法により評価指標を推定する第2推定手段、10は遺伝子の組み合わせと評価指標データの推定法を出力する出力手段である。

【0033】ここで、この発明の態様として入力には、手術で摘出された器官の癌患部試料と正常と判断された部位試料との各遺伝子の発現比率をDNAマイクロアレイで計測したデータ（発現プロファイルデータ）を用いた。また、被検試料の注目する評価指標として、その被

検試料を得た被験者の予後の存命月数を用いた。但し、存命月数が60ヶ月（5年）以上は該癌における影響はないものとして60ヶ月として取り扱う。第1推定手段構築手段4と第2推定手段構築手段8、第1推定手段5と第2推定手段9は同じ手段を用いてもよいし、違う手段を用いてもよい。

【0034】なお、上記の癌患部試料と正常部位試料における遺伝子発現プロファイルデータと評価指標データである存命月数は、この発明に使用される遺伝子発現プロファイルデータ、及び評価指標データのそれぞれの一例であり、この発明を何ら限定するものではないのは当然であり、目的の評価指標データを推定するために使用できる被検試料の遺伝子発現プロファイルデータ、評価指標データであればよい。

【0035】次に、この発明の実施の形態1に係る遺伝子発現プロファイル解析装置の使用方法について説明する。

【0036】使用方法の手順は、図8のフローチャートの通りであるので、図8を用いて説明する。まず、入力20ステップS1に各被検試料の遺伝子の発現プロファイルデータと評価指標データである予後の存命月数を入力する。

【0037】次に、データ前処理ステップS2で、遺伝子発現プロファイルデータの正規化、不必要的データの削除などを行う。この前処理は入力されたデータに基づき適切な方法を選択することができる。

【0038】次に、遺伝子組み合わせ候補生成ステップS3で遺伝子の組み合わせを生成する。この実施の形態1では、遺伝的アルゴリズムを用いて遺伝子の組み合わせを生成する。遺伝的アルゴリズムでは、複数の組み合

30 わせを生成する。遺伝的アルゴリズムでは、複数の組み合わせを生成し、その組み合わせの評価を行い、基準を満たす組み合わせが得られるまで探索を行う。基準を満たさない場合は、遺伝子の組み合わせの選択・交叉を行い、新たな遺伝子の組み合わせを生成し、評価・選択を行う。

【0039】遺伝子の組み合わせを表現する方法として、図2のようなビット列を用いた。DNAアレイで計測した各遺伝子に対応したビットを設け、そのビットが1である場合、その遺伝子を遺伝子の組み合わせに採用し、0である場合採用しない。このようなビット列を複数個用意し（ビット列のセット）、遺伝的アルゴリズムでビット列の生成・評価を行い、注目する評価指標データ（この実施の形態1では予後の存命月数）を推定するのに適切な遺伝子の組み合わせを探索する。

【0040】第1回目のビット列のセットの生成では、ランダムに各ビット列を生成する。2回目以降のビット列のセットの生成では、前回のビット列のセットを基に次のような方法で、ビット列のセットを生成する。ビット列に対する評価が最も高いいくつかのビット列はそのまま残す。残りのビット列は、前回のビット列のセット

11

の中から任意の2つのビット列を選択し、交叉によって新たなビット列を生成する。ここでの交叉は、各ビット毎にどちらのビット列のビットを選択するか決める一様交叉を用いる。交叉によって生成したビット列にさらに突然変異を加えて、新たなビット列を生成する。ここで、各ビットをランダムに反転させるような突然変異を*

を0に変える確率 λ_{out} と0を1に変える確率 λ_{in} とに分けて突然

変異率を設定するようにした。この2つの確率間に数1

を満たすような関係を導入することによって、1の数を10大きく変化させないようにできる。

【0042】

※

ここで、 G_{in} は採用する各遺伝子の発現比率、 G_{out} は予後の存

命月数、 α は係数を表す。

【0043】次に、第1推定方法生成ステップS4で各遺伝子の組み合わせ毎に推定方法を構築する。推定方法として、採用した遺伝子の重み付き線形和による推定である数2を採用した。

ここで、 x_i は採用する各遺伝子の発現比率、 y は予後の存命月数、 α_i は係数を表す。

【0045】次に、第1推定ステップS5で予後の存命月数を推定する。最小2乗法により予後の存命月数を推定するための係数を推定する。

【0046】次に、遺伝子組み合わせ評価ステップS6で各遺伝子の組み合わせの評価を行う。推定方法S5で★

存命月数の予測値と実際の値との平均誤差 \bar{e} に基づく評価

は数3で行った。

【0048】

【数3】

ここで、 f_{count} は平均誤差 \bar{e} に基づく評価、 a 、 b 、 c は定数である。図3は数3のグラフの一例を示す。図3のような評価関数により、望ましくない遺伝子の組み合せを効果的に排除し、有望な組み合わせを幅広く受け入*

また、採用している遺伝子数 $|G_{in}|$ に関する評価は数4で行った。

【0050】

$$f_{count} = 0$$

※※【数4】

if $|G_{in}| = 0$ or $n_{sample} \leq |G_{in}|$

$$= 1 - (1 - s) \frac{|G_{in}|}{\hat{G}_{in}} \quad \text{if } 0 < |G_{in}| < \hat{G}_{in}$$

$$= s \frac{n_{sample} - |G_{in}|}{n_{sample} - \hat{G}_{in}} \quad \text{if } \hat{G}_{in} \leq |G_{in}| < n_{sample}$$

ここで、 f_{count} は遺伝子数 $|G_{in}|$ に関する評価、 n_{sample} は全被検試料数、 s 、 \hat{G}_{in} は定数である。

12

*行うと、1の数が約半数になるように変化してしまう。

【0041】この発明の課題では、遺伝子の採用数は入力された遺伝子数の半数より非常に少くないと推定の汎用性が小さいと考えられるので、採用数が多いビット列は評価が低い。そこで、突然変異の前後でビット列における1の数が大きく変化しないように、1

$$\lambda_{out} |G_{in}| = \lambda_{in} |G_{out}|$$

$$y = \sum_i \alpha_i x_i$$

★【0044】

【数2】

★求めた推定法によって推定した予後の存命月数の誤差と

採用した遺伝子数の2つの観点で遺伝子の組み合わせを

評価した。

【0047】

◆

◆求めた推定法によって推定した予後の存命月数の誤差と

採用した遺伝子数の2つの観点で遺伝子の組み合わせを

評価した。

【0049】

★求めることが可能になっている。

【0051】図4は、数4の説明図である。数4は、遺伝子の数に対して被検試料の数が圧倒的に少ないため採用する遺伝子数が多くなると推定に汎用性がなくなってしまうという問題は、少なくとも採用遺伝子数が予め設定する \hat{G}_{in}

より小さくすることにより解決されることを見出し、こ※※の発明においては \hat{G}_{in} を被検試料数未満、好ましくは充分に小さくすることにより解決した。

【0052】数3、数4の評価を組み合わせて各遺伝子★せは、数5のように f_{error} と f_{count} との加重平均により決定する。こうすることにより α の調整により推定誤差と採用遺伝子数

とのバランスを図ることができ、被検試料の数が極端に少なく推定誤差がある程度犠牲にする必要がある場合などに對話的に遺伝子を抽出することができる。

【0053】☆ 次に、遺伝子組み合わせ抽出ステップS7で、数5で求めた f を用いて遺伝子組み合わせ群を抽出する。遺伝子組み合◆◆わせ

群は f が十分大きい値となつた遺伝子組み合わせの集合である。

【0054】次に、遺伝子組み合わせ制約ステップS8で、抽出された遺伝子組み合わせ群の中で制約条件を満たしているものを抽出する。遺伝子の機能や遺伝子同士の関連などに関する情報を利用できるならば、遺伝子組み合わせ制約ステップS8でその情報を入力し、遺伝子*

合わせを表現すると、通常 n 個の遺伝子に対応するビット列は長さ n のビット列として表現され 2^n 通りの組み合わせが考えられるが、既知の知識からその内 $p < 2^n$ 通りの組み合わせを考慮するだけでいいことがわかっている場合、長さ n のビット列の代わりに $\log_2 p$ 以上で最小の整数だけの長さのビ

ト列を用いる。こうすることにより、遺伝的アルゴリズムの探索空間を小さくすることができると同時に、既知の知識を推定に役立つ部分ビット列として効果的に利用することができる。

【0055】遺伝子組み合わせ評価ステップS6で求めた各遺伝子の組み合わせの評価を比較し、最もよい評価があらかじめ設定した値よりよければ、遺伝子の組み合わせの選択を終了し、第2推定方法生成ステップS9に移る。設定した値以上の遺伝子の組み合わせがなければ、遺伝子組み合わせ候補生成ステップS3に戻って遺伝子の組み合わせの生成・推定・評価・選択を繰り返す。

【0056】次に、第2推定方法生成ステップS9では、求めた最適な遺伝子の組み合わせを用いて予後の存命月数を推定する手段を構築する。第1推定方法生成ステップS4で用いた線形多項式を採用してもよいし、誤差逆伝播学習を用いたニューラルネットワークを採用し

*用する遺伝子数が多くなると推定に汎用性がなくなってしまう

★の組み合わせを評価する。組み合

☆【数5】

$$f = \alpha f_{\text{error}} + (1 - \alpha) f_{\text{count}}$$

*の組み合わせにおける制約条件を生成する。既知の知識は、複数遺伝子の同時利用、同時利用禁止などの制約として表現する。すなわち、遺伝子組み合わせ候補生成ステップS3で説明した図2のビット列で遺伝子の組み

てもよい。ニューラルネットワークを用いる場合、予め相間の高い遺伝子の組み合わせに絞り込んだ上で学習することができるため、被検試料データが少ないと起因する過学習を防ぐことができる期待される。

【0057】次に、第2推定ステップS10では、採用した推定手段を用いて予後の存命月数を推定する。線形多項式の場合には、最小2乗法を用いればよいし、ニューラルネットワークの場合には誤差逆伝播学習則を用いて係数を学習する。

【0058】最後に出力ステップS11から、求めた最適な遺伝子の組み合わせ群と予後の存命月数の推定法を出力する。すなわち、求めた遺伝子の発現プロファイルデータから、求めた推定法により予後の存命月数を推定することができる。

【0059】図5は、予後の存命月数との相関係数の大きい順に選んだ25個の遺伝子を用いて予後の存命月数を線形多項式を用いて推定した結果である。予後の存命

40

50

月数の誤差の平均値は10.7ヶ月であり、よい推定はできなかった。

【0060】図6は、この発明で求めた最適な遺伝子の組み合わせの一例で推定した結果である。予後の存命月数の誤差の平均値は2.6ヶ月であり、精度のよい推定ができた。

【0061】この発明により一度求めた、最適な遺伝子の組み合わせ群と予後の存命月数の推定法をデータベースに蓄積することにより、入力される遺伝子発現プロファイルを時間の経過とともに増加させることができ、例えば、第2推定方法生成ステップS9の学習効果を高めることができ、データの蓄積が進めば、少量のデータで精度の高い評価指標の推定値を得ることが可能になる。

【0062】また、一度データベースに蓄積した遺伝子組み合わせ群と存命月数に代表される評価指標の推定法のデータの再利用方法として、特に限定するものではないが、例えば、小型のパーソナルコンピュータにインストールし、ヒトを含む生物の生体内環境、例えば分化、成長、老化、代謝のモニタリング、疾病の診断等が挙げられる。

【0063】上記のような、この発明により選択された遺伝子の発現プロファイルを測定する態様としては、特に限定するものではないが、例えば、被検試料中で発現されている前記遺伝子由来のmRNA量の測定が挙げられる。mRNA量の測定方法には、適切なプローブを使用するハイブリダイゼーション法やRT-PCR法に代表される核酸增幅法を利用することができますが、これらに限定されるものではなく、公知のmRNA定量方法のいずれもが使用できる。同時に複数のmRNA種の定量を行う観点からは、前記遺伝子又はその断片が担体上に固定化されたDNAアレイを使用したハイブリダイゼーション法が好適である。ここでDNAアレイとは、1以上のDNA分子種がそれぞれ担体上のあらかじめ定められた領域に固定化されたものを称する。当該DNAアレイに使用される単体には特に限定ではなく、例えば、平板上、膜状、テープ又はひも状、粒子状の担体が使用できる。粒子状単体を使用したアレイとしては、個々のビーズに特定のDNA分子種が固定化されたマイクロビーズアレイが例示される。また、固定化されたDNAと試料中の核酸とのハイブリッドの形成を直接検出する手段、例えばハイブリッド形成の有無を電気信号として出力する手段を備えたDNAアレイであってもよい。

【0064】さらに、この発明は、上記のように選択された遺伝子の発現プロファイルの測定に有用なDNAアレイ並びにmRNA量測定用のキットを提供する。この発明のDNAアレイは、上記のようにこの発明により選択された遺伝子、又はその断片が担体上のあらかじめ定められた領域に固定化されたものである。また、この発明のキットは選択された遺伝子から転写されるmRNAをそれぞれ定量可能なものであり、この発明により選択

された遺伝子、又はその断片にハイブリダイズ可能なプローブを含むキット、並びに前記遺伝子、又はその断片を核酸增幅法によって増幅するための少なくとも1種のプライマーを含むキットが例示される。これらのキットは、ハイブリダイゼーション、又は核酸増幅方法を実施するための試薬類をさらに含むものが好適である。

【0065】試料に関する評価指標データを得るための装置もこの発明により提供される。該装置は下記手段：

(1) この発明の遺伝子発現プロファイル解析方法を用いて選択された遺伝子の発現プロファイルデータを入力する手段、(2) この発明の遺伝子発現プロファイル解析方法を用いて提供される評価指標データの推定法により、手段(1)で得られた遺伝子発現プロファイルから被検試料における評価指標データを推定する手段、及び(3)手段(2)で得られた評価指標データを出力する手段。

【0066】を備えていることを特徴とする。

【0067】ここで、手段(1)としては、さらに前記の方法を用いて選択された遺伝子の発現を測定する手段を備えていてもよく、例えば前記この発明により選択された遺伝子、及びその断片が担体上の予め定められた領域に固定化されたDNAアレイを備えたものが例示される。この場合、DNAアレイ上でのハイブリダイゼーションを実施するための装置、試料由来の核酸とハイブリダイズしたDNAを検出するための装置をさらに備えたものが好ましい。手段(3)としては、評価指標データをディスプレイ上、印刷物、及び/又は磁気若しくはその他の記憶手段を備えた記憶媒体上に出力可能なもののが挙げられる。

【0068】この発明は、以上説明したように構成されているので、以下に示すような効果を奏する。

【0069】データを入力する入力手段と、遺伝子の組み合わせを生成する遺伝子組み合わせ手段と、評価指標データを推定するための手段を構築する第1推定手段構築手段と、評価指標データを推定する第1推定手段と、生成した遺伝子の組み合わせを評価する遺伝子組み合わせ評価手段と、遺伝子の組み合わせと推定法を出力する出力手段とを備えることにより、適切な遺伝子の組み合わせを選択することができ、注目する評価指標データを推定する推定法を得ることができる。

【0070】また、不必要的遺伝子を除いたり、データの正規化を行うデータ前処理手段によりデータの前処理を行うことにより、効率よく遺伝子の組み合わせの生成・推定・評価を行うことができる。

【0071】また、求めた最適の遺伝子の組み合わせを用いて評価指標データを推定する第2推定手段構築手段と第2推定手段を備えることにより、よりよい推定を得ることができる。

【0072】また、遺伝子組み合わせ手段における組み合わせを制約する組み合わせ制約手段を備えたことによ

り、既存の遺伝子の機能や遺伝子同士の関連などに対する知識を利用することができます、より必然性の高い推定法を得ることができる。

【0073】また、遺伝子組み合わせ手段に遺伝的アルゴリズムを用いることにより、効率よく最適な遺伝子の探索を行うことができる。

【0074】また、遺伝子組み合わせ手段に用いる遺伝的アルゴリズムの評価に、評価指標データの推定誤差と採用する遺伝子数に基づく評価を用いることにより、推定誤差が小さく汎用性のある推定を得ることができる。

【0075】また、遺伝子組み合わせ手段に用いる遺伝的アルゴリズムの遺伝子の組み合わせの変化法として、遺伝子の採用と非採用を同程度変化させるようにしたので、不必要的変化を減少させ、効率よく最適な遺伝子の組み合わせを探索できる。

【0076】また、第2推定手段構築手段として、誤差逆伝播学習を用いたニューラルネットを用いたので、より精度の高い推定法を得ることができる。

【0077】この発明の遺伝子発現プロファイル解析装置を使用して、目的とする評価指標データを得るために最適な遺伝子の組み合わせを探索し、同時に、選択された遺伝子の発現プロファイルデータから個々の試料における評価指標データを推定するための推定方法を構築することができる。この発明を利用する態様としては、特に限定するものではないが、例えばヒトを含む生物の生体内環境、例えば分化、成長、老化、代謝のモニタリング、疾病の診断等が挙げられる。

【0078】例えば、ヒトの疾病的診断や治療法の選択においては、当該疾病的特徴を分子レベルで把握することが有用である。各種の疾病には多数の遺伝子の発現変動が起こっているが、当該変動はその原因、病状、或いは個体差によってそのパターンが異なる。すなわち、いくつかの遺伝子の発現プロファイルデータは各個体における疾病的性格や病状を反映しており、その解析によって診断、治療に有用な評価指標に関するデータを得ることが可能である。前記評価指標としては、例えば、疾病的名称、タイプ、原因、進行状況、予後、余命、薬剤に対する感受性やその副作用、発症、再発、転移の可能性等がある。この発明により、これらの評価指標の推定に有用な遺伝子の選択、並びに評価指標データの推定法を効率よく、かつ高い精度で得ることができる。また、疾病を有していない個体においても、適切に選択された遺伝子の発現プロファイルから健康状態や各種疾病的発症の可能性を評価指標として得ることができる。

【0079】このような解析が有用な疾患としては、特に限定するものではないが、例えば、癌、糖尿病、肝疾患、循環器系疾患（心筋梗塞、脳血管障害等）、高血圧、遺伝性疾患、感染症等が挙げられる。また、寿命、肥満、禿頭等の評価指標データを得ることもできる。

【0080】上記の他、生理活性物質の探索や薬物代謝

の研究など、生物学、分子生物学、医学、薬学領域における基礎研究にもこの発明を利用することが可能である。

【0081】

【実施例】以下、この発明を実施例をもって更に具体的に説明するが、この発明はこれらの実施例によって限定されるものではない。

【0082】【実施例1】癌関連遺伝子搭載cDNAチップの作製と胃癌検体における遺伝子発現プロファイルの解析

（1）胃の癌化により発現が変化する遺伝子由来cDNAの単離
胃の癌化病変部試料と対照正常部試料とを比較するディファレンシャルディスプレイ法を行い、癌化により発現量の変化する遺伝子由来cDNAを単離した。

【0083】まず、インフォームド・コンセントの得られた低分化腺癌の進行癌患者から摘出された胃の癌化病変部及び対照正常部のそれぞれから、TRIZolTM試薬（ギブコBRL社製）を用いて、全RNAを抽出し、粗RNA試料とした。得られた粗RNA試料のうち50μg相当量と、最終濃度5mM MgCl₂と20単位のRNase阻害剤（宝酒造社製）と10単位のDNase I（宝酒造社製）とを、37°C、30分間反応させ、ゲノムDNAを除去し、RNA試料を得た。

【0084】前記RNA試料について、Fluorescence Differential DisplayTM Kit Rhodamine version（宝酒造社製）とEnzyme Set-FDD（宝酒造社製）とを用いて、キット添付の説明書記載の手順に従い、下記の通りにRT-PCRを行った。

【0085】まず、前記RNA試料200ngと、前記Fluorescence Differential DisplayTM Kit Rhodamine versionに添付された、5'-CG-3'のアンカーリ配列を有するローダミンラベル下流プライマーとを混合して得た混合液を70°C、10分間の熱処理後、急冷した。この混合液にAMV逆転写酵素とdNTPとを添加して、55°C、30分の逆転写反応を行い、一本鎖cDNA試料を得た。

【0086】前記一本鎖cDNA試料を錆型とし、逆転写反応時と同じローダミンラベル下流プライマーと、前記キットの24種の上流プライマー（R1乃至R24）のいずれか1種とを用いて、1. 3mM MgCl₂と、基質として各100μMのdATP、dGTP、dCTP及びdTTPとを含有した反応液中でPCRを行い、合計24種類の増幅DNA試料を得た。なお、前記PCRのサーマルプロファイルは、94°C、2分-40°C、5分-72°C、5分の処理のあと、94°C、30秒-40°C、2分-72°C、1分からなる反応を1サイクルとして34サイクル、その後、さらに72°C、5分

の処理、とした。

【0087】反応終了後の反応液に等量の9.5%ホルムアミドを添加し、90℃、2分間熱変性して電気泳動用*

光イメージアナライザーフォトビオII

で読み取り、検出した。その結果、多数のバンドからなるフィンガープリントが得られ、癌病変部由来試料と対照正常部由来試料とで強度の異なる複数のバンドが存在した。

【0088】ついで、上記のフィンガープリント上に、前記電気泳動後のポリアクリルアミドゲル板を置き、癌病変部と対照正常部との間で強度の異なるバンド部分のゲル250本を切り出した。ゲルからDNA断片を水抽出し、Cloning- Sequencing Primer Set for FDD(宝酒造社製)中の各DNAフラグメントに対応するプライマーを用い、再度PCRによる增幅を行った。

【0089】増幅されたフラグメントをTAクローニング法にてクローニングし、1フラグメントにつき4クロー

10

*サンプルとした。このサンプルを7M尿素変性4%ポリアクリルアミドゲル電気泳動に供し、得られたゲル上のDNA断片泳動パターンを蛍

Mult i - View(宝酒造社製)

ンズつ単離した。ついで、得られた各クローンに含まれる核酸について、ダイテオキシ法により、塩基配列を決定した。得られた塩基配列をParacel Clustering packageソフト(Paracel社製)を用いてアセンブリを行い、形成された各コンティグ(contig)配列について、塩基配列情報を収録したデータベースを用いたホモロジー探索を行った。

【0090】以上のようにして得られた遺伝子のうち、65遺伝子は既に単離、同定された遺伝子であるが、その癌との関連は知られていないことが明らかとなった。これらの遺伝子名並びにGenBankにおける登録番号(Accession No.)を表1に示す。

【0091】

【表1】

No.	遺伝子名	Accession No.
1	H.sapiens mRNA for RING protein	Y07828
2	Homo sapiens mRNA for PKU-beta, com	AB004885
3	Homo sapiens carboxy terminus of Hsp70-interacting protein (CHIP) mRNA, complete cds	AF129085
4	Homo sapiens mRNA; cDNA DKFZp586I0523 (from clone DKFZp586I0523)	AL050217
5	H.sapiens Sp17 gene	Z48570
6	Homo sapiens MD-1 mRNA, complete cds	AF057178
7	H.sapiens mRNA for BiP protein	X87949
8	Human chromosome 17q12-21 mRNA, clone pOV-2, partial cds	U18919
9	Homo sapiens dead box, X isoform (DBX) mRNA, alternative transcript 2, complete cds	AF000982
10	Human ferritin heavy chain mRNA, complete cds	M97164
11	Human ovarian cancer downregulated myosin heavy chain homolog (Doc1) mRNA, complete cds	U53445
12	Homo sapiens full length insert cDNA clone YP42A04	AF085884
13	Human mRNA for protein disulfide isomerase-related protein P5, complete cds	D49489
14	Homo sapiens Ku70-binding protein (KUB3) mRNA, partial cds	AF078164
15	Human nonmuscle myosin heavy chain-B (MYH10) mRNA, partial cds	M69181
16	Human voltage-dependent anion channel isoform 1 (VDAC) mRNA, complete cds	L06132
17	Homo sapiens phosphatidylinositol 4-kinase mRNA, complete cds	L36151
18	Human beta adaptin mRNA, complete cds	M34175
19	Homo sapiens H beta 58 homolog mRNA, complete cds	AF054179
20	Homo sapiens unknown mRNA, complete cds	AF047439
21	Homo sapiens NRD convertase mRNA, complete cds	U64898

22	H.sapiens mRNA for acylphosphatase, muscle type (MT) isoenzyme	X84195
23	Homo sapiens transmembrane protein BRI (BRI)	AF152462
24	Homo sapiens dynactin subunit (p22) mRNA, complete cds	AF082513
25	Human mRNA for aldose reductase (EC 1.1.1.2)	X15414
26	Homo sapiens cDNA FLJ20693 fis, clone	AK000700
27	Homo sapiens mRNA for scrapie responsive protein 1o	AJ224677
28	Homo sapiens KIAA0402 mRNA, partial cds	AB007862
29	Human transactivator protein (CREB) mRNA, complete cds	M27691
30	Human MAL protein gene mRNA, complete cds	M15800
31	Homo sapiens ataxin-2-like protein A 2LP (A2LG) mRNA, complete cds	AF034373
32	Human phosphatidylinositol transfer protein mRNA, complete cds	M73704
33	H.sapiens mRNA for transketolase	X67688
34	Homo sapiens nibrin (NBS) mRNA, complete cds	AF051334
35	Human SnRNP core protein Sm D3 mRNA, complete cds	U15009
36	Homo sapiens FUS/TLS protein gene,	AF071213
37	Homo sapiens clone 486790 diphosphoinositol polyphosphate phosphohydrolase mRNA, complete cds	AF062529
38	Human globin gene	M69023
39	H.sapiens mRNA for aminopeptidase P-like	X95762
40	H.sapiens mRNA for UDP-GalNAc:polypeptide N-acetylgalactosaminyl transferase	X92689
41	Homo sapiens cDNA FLJ20463 fis, clone	AK000470
42	Novel gene similar to C. elegans hypothetical 55.2 KD protein F16A11.2, SW: P90838	AL050255
43	Homo sapiens cDNA FLJ10986 fis, clone	AK001848
44	Human inhibitor of apoptosis protein 2 mRNA, complete cds	U45879
45	Human MB-1 mRNA, complete cds	M80462
46	Human tissue factor gene, complete cds. 1/1995	J 02846

47	Homo sapiens mRNA; cDNA DKFZp566E0224 (from clone DKFZp566E0224)	AL050031
48	Homo sapiens NADH-ubiquinone oxidoreductase subunit CI-B12 mRNA, complete cds	AF047183
49	Homo sapiens ASH1 mRNA, complete cds. 5/2000	AF257305
50	Homo sapiens CGI-65 protein mRNA, complete cds	AF151823
51	Homo sapiens clone 24451 mRNA sequence	AF070599
52	Homo sapiens ubiquitin hydrolyzing enzyme I (UBH1) mRNA, partial cds	AF022789
53	Homo sapiens mRNA for cysteine-rich protein	AJ006591
54	Human mRNA for KIAA0139 gene, complete cds	D50929
55	Homo sapiens clone 23819 white protein homolog mRNA, partial cds	AF038175
56	Homo sapiens connexin 26 (GJB2) mRNA, complete cds	M86849
57	Homo sapiens mRNA for KIAA0914 protein, complete cds	AB020721
58	Homo sapiens mRNA for KIAA0907 protein, complete cds	AB020714
59	H. sapiens mRNA for transcriptional intermediary factor 2	X97674
60	Human cytosolic aspartate aminotransferase mRNA, complete cds	M37400
61	Homo sapiens CGI-118 protein mRNA, complete cds	AF151876
62	Human guanylate binding protein isoform II (GBP-2) mRNA, complete cds	M55543
63	Homo sapiens cDNA FLJ10633 fis, clone	AK001495
64	Homo sapiens chaperonin containing t-complex polypeptide 1, eta subunit (Ccth) mRNA, complete cds	AF026292
65	Homo sapiens mRNA for Prer protein	AJ005579

さらに、上記の遺伝子の他、データベースを用いたホモロジー検索では既知の遺伝子とのホモロジーが見出されなかった171種の遺伝子断片が得られた。これらの遺伝子断片の塩基配列をそれぞれ配列表の配列番号1乃至171に示す。これらの各群の遺伝子の組み合わせは、いずれも平均誤差2.0~3.0ヶ月で存命月数を推測することができた。

【0092】(2) DNAアレイの作製

前記(1)記載の、胃の癌化により発現変動することが見出された各遺伝子由来増幅DNA断片を用い、DNAアレイ搭載フラグメントを調製した。

【0093】まず、(1)記載の方法でDD法により見出されたクローンのうち、各コンティグを代表する配列を含有したクローン、並びにWO98/37187号国際公開パンフレットに記載された癌関連遺伝子のクローンの計346種類のクローンを鋳型とし、当該プラスミ

ドのマルチクローニングサイトの両端に設定されたプライマーを使用したPCR法により目的のcDNA断片を増幅した。ついで、増幅されたcDNA断片の塩基配列分析を行って、目的の断片であることを確認した。また、目的の断片をエタノール沈殿法により回収し、回収された断片を100mM炭酸緩衝液(pH9.5)に1μMとなるように溶解した。

【0094】この他、ハウスキーピング遺伝子としてβ-アクチン遺伝子、チューブリンα2、シクロフィリン、グリセルアルデヒド3磷酸デヒドロゲナーゼ、リボゾーマルタンパク質S5、リボゾーマルタンパク質S9等の遺伝子を、またネガティブ対照として、プラスミドpUC19をそれぞれ同様に調製した。これらをDNAチップ作製装置(アフィメトリクス社製)を用いて、アミノ基導入スライドガラス(シグマ社製)にスポットし、UV照射により固定した。スライドを0.2%SD

S、次いで蒸留水で洗浄乾燥してDNAアレイとした。
【0095】(3)蛍光標識cDNAの調製
遺伝子発現解析には、インフォームド・コンセントの得られた35名の胃癌患者由来の試料を用いた。これら35名の胃癌患者の臨床病理因子の内訳を以下に示す。

【0096】①進行度：

進行度1（癌の浸潤程度が浅く、転移の認められない早期の癌）：7名

進行度2（癌が浸潤、近接リンパ節に転移した状態）：8名

進行度3（遠隔リンパ節に広がった状態）：4名

進行度4（より遠隔リンパ節に広がるとともに腹膜転移、肝転移した状態）：16名

②分化度：

乳頭腺癌（pap）：1名

高分化腺癌（tub）：4名

中分化腺癌（tub2）：11名

低分化腺癌（por）：15名

印環細胞癌（sig）：2名

粘液癌（muc）：2名

③深達度（癌浸潤の及んだ最も深い胃壁の層で表す）：

粘膜（m）：2名

粘膜下層（sm）：4名

筋層（mp）：5名

しょう膜化組織（ss）若しくはしょう膜を破り腹膜に露出（se）：24名

④リンパ節転移

なし：8名

あり：27名

⑤腹膜播種

なし：32名

あり：3名

それぞれの胃癌患者より癌摘出手術等に摘出された組織から胃癌組織と対照正常胃組織を取り分けた。ついで、AGPC (Acid Guanidium Phenol-Chloroform) 法により、各組織から個々に全RNAを摘出した。これらの全RNAのそれぞれからOligotex-MAG mRNA Purification Kit (宝酒造社製) を用いてmRNAを射精した。

【0097】上記mRNAを錆型とし、逆転写酵素を用いてcDNA合成反応を行った。なお、対照正常胃組織群の場合、Cy3-dUTP (アマシャム社製) を含むdNTPを用い、胃癌組織群の場合、Cy5-dUTP (アマシャム社製) を含むdNTPを用いた。なお、各検体由来胃癌組織と対照正常胃組織につき2反応ずつ標識反応を行った。反応液組織を以下に示す。

【0098】反応液A：上記mRNA約1μg、300pmolのオリゴdTプライマー (宝酒造社製) 及び最終的に11.9μlになるようにジエチルピロカーボネ

ート (DEPC、ナカライトスク社製) 処理水を添加。

【0099】反応液B：5×AMV RTase用緩衝液 (ライフサイエンス社製) 4μlと、各0.1mMのdATP、dCTP、dGTP及び0.065mMのdTTPと、30UのRNaseインヒビター (宝酒造社製) と、0.035mMのCy3又はCy5標識dUTP (アマシャムファルマシア社製) とを混合し、最終容量6.5μlの溶液を得た。

【0100】反応液Aを70℃で10分間保持した後、

10 氷浴上で冷却し、反応産物を得た。ついで、前記反応産物に、反応液Bと約30単位のAMV RTase (ライフサイエンス社製) とを添加し、さらに55℃で30分間保持した。その後、得られた反応産物に約30単位のAMV RTaseをさらに添加し、反応液量を20μlにし、RT反応液を得た。前記RT反応液を42℃で60分間保持した。ついで、この反応液を70℃で10分間保持し、逆転写反応を停止し、室温まで冷却した。得られた反応産物を、Centri-sep spin Column (アプライド・バイオシステムズ社製) を用いてゲルろ過した。このようにして得られたCy3標識cDNA、Cy5cDNAを、同一患者分を混合してエタノール沈殿濃縮し、ハイブリダイゼーション緩衝液 (6×SSC/0.2%SDS/5×デンhardt液/0.1mg/mlサケ精子DNA) 10μlに溶解して、蛍光標識cDNAを調整した。なお、1×SSCの組織は、0.15M NaCl、0.015Mクエン酸ナトリウム、pH7.0である。

【0101】(4)ハイブリダイゼーション

前記(2)で作製したDNAアレイ、及び市販のTakara IntelliGene™ Human Cancer CHIP version 2.0 (ヒト由来既知遺伝子のうち癌疾患に関連する遺伝子425種類搭載、宝酒造社製) に、ブレハイブリダイゼーション緩衝液 (6×SSC、0.2%SDS、5×デンhardt液、1mg/mlサケ精子DNA) を滴下し、カバーグラスをかけて室温で2時間保持した。その後、カバーグラスを除き、アレイを2×SSCで洗浄し、ついで、0.2×SSCで洗浄し、風乾した。

【0102】ついで、前記(3)で調製した標識cDNA各2本を熱変性した後、1反応分全量を(2)で調製したDNAアレイに、残り1反応分全量をTakara IntelliGene (商標) Human Cancer CHIP version 2.0に滴下し、カバーグラスをかけて周囲をフィルムで密閉した。これを65℃で16時間保持した。ついで、カバーグラスを除いて、アレイを、0.2×SSC/0.1%SDS中、55℃で30分の洗浄を2回行い、ついで、65℃で5分洗浄し、さらに、0.05×SSC中、室温で5分洗浄し、風乾した。

50 【0103】風乾後のアレイを、マイクロアレイスキ

ナー（アフィメトリクス社製）にかけて各スポットの蛍光シグナルを解析した。測定されたシグナルを発現データ解析ソフト *Image* (バイオディスカバリー社製) で解析し、個々の胃癌患者について、癌組織と対照正常組織における各遺伝子の発現量を調べた。

【0104】(5) 遺伝子発現データの解析
各遺伝子の発現変動を、癌組織の対照正常組織に対する各遺伝子の相対発現比率【癌組織の発現強度シグナル／対照正常組織の発現強度シグナル】により判定した。この際、比較対照となる癌組織mRNAと対象正常組織mRNAの質の補正のため、ハウスキーピング遺伝子の発現量の平均値で補正し、ノーマライゼーションを行った。この結果、35名の胃癌患者における、前記(2)で作製したDNAアレイ、及びTakara IntelliGene™ Human Cancer CHIP version 2.0 に搭載された合計771種類の遺伝子発現プロファイルを得た。

【0105】前記771種類の遺伝子の35名の胃癌患者における発現比率と予後の存命月数との相関関係を求めた。そのヒストグラムを図7に示す。

【0106】【実施例2】遺伝子発現プロファイル解析

(1) 周知の方法を用いた予後の存命月数の推定
 評価指標を予後の存命月数とし、35名の被験者由來の被検試料より前述実施例1-(5)で得られた771種類の遺伝子発現プロファイルデータと予後の存命月数を用いた。図7に示したヒストグラムの中から、相関係数の大きい順に選んだ25個の遺伝子を用いて予後の存命月数を線形多項式を用いて推定した結果を図5に示す。予後の存命月数の誤差の平均値は10.7ヶ月であり、上位推定はできなかった。

【0107】(2) この発明の方法を用いた予後の存命日数の推定

評価指標を予後の存命月数とし、35名の被験者由来の被検試料より前述実施例1-(5)で得られた771種類の遺伝子発現プロファイルデータと予後の存命月数を入力データとする。存命月数が60ヶ月(5年)以上は該癌に影響はないものとして60ヶ月として扱う。入力された遺伝子発現プロファイルデータは、前処理として、発現比率データの正規化、不要データの削除を行った後、各遺伝子を1ビットとして長さ771のビット列を生成する。遺伝的アルゴリズムで500個のビット列の生成を行った。1回目の生成では、ランダムに500個のビット列を生成した。2回目以降の生成では、最も評価が高かったビット列を1つ残し、残りのビット列は前回のビット列のセットから任意の2つのビット列を選択し、交叉によって新たなビット列を作り、それに突然変異を加えることによって生成した。交叉は一様交叉とし、突然変異は、突然変異率の数1を用いて、入

で設定した。ここで、nは値が1のビットの数を表す。実測値と予測値との平均誤差と採用した遺伝子数により、各遺伝子組み合わせを評価した。平均誤差評価は、数3を使用し、それぞれa = 0. 5 2 5 1 8 6 7 4 7 5、b = 16. 5 0 2 7 2 5 6 6、c = 0. 3 7 6 5 5 9 0 6 2 5とした。また、採用遺伝子数の評価は数4でs = 0. 9、G_{1..n} = 25である。

〔0108〕これらに対して重み付けを行い、最終的に数5で表される評価関数、 $\alpha = 0.9$ で評価を行った。

10 [0109] 評価を繰り返し行い、下記に第1群として示した24種の遺伝子の組み合わせが得られた。求まつた最適な遺伝子の組み合わせで予後推定をした結果を図6に示す。誤差平均が2.6ヶ月となり、精度のよい推定ができた。

〔0110〕1群(24種)

- "retinoic acid receptor,beta" (GenBank Accession No. Y00291)
- envoplakin (GenBank Accession No. U53786)
- insulin-like growth factor binding protein (GenBank Accession No. M35878)
- proliferating cell nuclear antigen (GenBank Accession No. M15796)
- platelet-derived growth factor beta polypeptide (simian sarcoma viral(v-sis)oncogene homolog) (GenBank Accession No. M12783)
- Homo sapiens mRNA for Prer protein (GenBank Accession No. AJ005579)
- Human CO-029 (GenBank Accession No. M35252)
- 配列表の配列番号 1 5 9 に記載された塩基配列を有する遺伝子
- "collagen,type VIII, alpha 1" (GenBank Accession No. X57527)
- frizzled-related protein (GenBank Accession No. U91903)
- Human cell adhesion protein (vitronectin) receptor alpha subunit mRNA, complete cds" (GenBank Accession No. M14648)
- GC36 (WO 98/37187号国際公開パンフレット)
- 配列表の配列番号 8 4 に記載された塩基配列を有する遺伝子
- 配列表の配列番号 9 5 に記載された塩基配列を有する遺伝子
- ectodermal-neural cortex (with BTB-like domain) (GenBank Accession No. AF059611)
- Ric (Drosophila)-like, expressed in many tissues (GenBank Accession No. Y07566)
- "Homo sapiens MD-1, RP105-associated (MD-1), mRNA A" (GenBank Accession No. NM_004271)
- "Homo sapiens cDNA FLJ10986 fis, clone" (GenBank Accession No. NM_004271)

Accession No. AK001848)

- ・配列表の配列番号 160 に記載された塩基配列を有する遺伝子
- ・配列表の配列番号 101 に記載された塩基配列を有する遺伝子
- ・"tumor necrosis factor receptor superfamily, member 18" (GenBank Accession No. M32315)
- ・配列表の配列番号 16 に記載された塩基配列を有する遺伝子
- ・配列表の配列番号 56 に記載された塩基配列を有する遺伝子
- ・"integrinalpha L (antigen CD11A (p180), lymphocyte-associated antigen 1; alpha polypeptide)" (GenBank Accession No. AC002310)

さらに、上記の第1群の遺伝子の選択とは独立した遺伝子発現プロファイル解析を実施し、下記に第2群～第8群として示した、7通りの遺伝子の組み合わせが得られた。これらの各群の遺伝子の組み合わせは、いずれも平均誤差2.0～3.0ヶ月で存命月数を推測することができた。

【0111】第2群（25種）

- ・配列表の配列番号 6 に記載された塩基配列を有する遺伝子
- ・配列表の配列番号 159 に記載された塩基配列を有する遺伝子
- ・"caspase 6, apoptosis-related cysteine protease" (GenBank Accession No. U20536)
- ・forkhead box C1 (GenBank Accession No. AF04869)

3)

- ・chondroitin sulfate proteoglycan 2 (versican) (GenBank Accession No. U16306)
- ・cell division cycle 10 (homologous to CDC10 of *S. cerevisiae*) (GenBank Accession No. S72008)
- ・配列表の配列番号 161 に記載された塩基配列を有する遺伝子
- ・v-Ki-ras2 Kirsten rat sarcoma 2 viral oncogene homolog (GenBank Accession No. M54968)
- ・cyclin D2 (GenBank Accession No. D13639)
- ・vascular endothelial growth factor B (GenBank Accession No. U43368)
- ・配列表の配列番号 1 に記載された塩基配列を有する遺伝子
- ・connective tissue growth factor (GenBank Accession No. X78947)

- ・cadherin 18 (GenBank Accession No. U59325)
- ・Human mRNA for SB classII histocompatibility antigen alpha-chain (GenBank Accession No. X03100)
- ・caveolin 2 (GenBank Accession No. AF035752)
- ・EphA2 (GenBank Accession No. M59371)
- ・"Homo sapiens mRNA for PKU-beta, com" (GenBank Accession No. X03100)

Accession No. AB004885)

- ・"Human tissue factor gene, complete cds. 1/1995" (GenBank Accession No. J 02846)
- ・配列表の配列番号 162 に記載された塩基配列を有する遺伝子
- ・配列表の配列番号 94 に記載された塩基配列を有する遺伝子
- ・配列表の配列番号 163 に記載された塩基配列を有する遺伝子
- ・"guanine nucleotide binding protein (G protein), alpha activating activity polypeptide, olfactory type" (GenBank Accession No. U55184)
- ・kinase insert domain receptor (a type III receptor tyrosine kinase) (GenBank Accession No. AF035121)
- ・O-6-methylguanine-DNA methyltransferase (GenBank Accession No. M29971)
- ・E2F transcription factor 3 (GenBank Accession No. D38550)

- ・"retinoic acid receptor, beta" (GenBank Accession No. Y00291)
- ・EphB2 (GenBank Accession No. AF025304)
- ・Homo sapiens full length insert cDNA clone YP42A04 (GenBank Accession No. AF085884)
- ・配列表の配列番号 164 に記載された塩基配列を有する遺伝子
- ・GG24 (WO 98/37187号国際公開パンフレット)
- ・配列表の配列番号 122 に記載された塩基配列を有する遺伝子
- ・interleukin 8 (GenBank Accession No. M26383)
- ・wingless-type MMTV integration site family member 2 (GenBank Accession No. X07876)
- ・apoptotic protease activating factor (GenBank Accession No. AF013263)
- ・"laminin, alpha 4" (GenBank Accession No. S78569)
- ・DNA-damage-inducible transcript 3 (GenBank Accession No. X92120)

- ・"pleckstrin homology, Sec7 and coiled/coil domains 2-like" (GenBank Accession No. U59752)
- ・配列表の配列番号 3 に記載された塩基配列を有する遺伝子
- ・配列表の配列番号 34 に記載された塩基配列を有する遺伝子
- ・keratin 19 (GenBank Accession No. Y00503)
- ・"tumor necrosis factor receptor superfamily, member 5" (GenBank Accession No. X60592)
- ・thyroid autoantigen 70kD (Ku antigen) (GenBank Accession No. X60592)

ccession No. Z83840)

- "integrin, beta 8" (GenBank Accession No. M73780)

· 配列表の配列番号 1 6 5 に記載された塩基配列を有する遺伝子

· 配列表の配列番号 4 1 に記載された塩基配列を有する遺伝子

- "collagen, type XI, alpha 2" (GenBank Accession No. AL031228)

· "Homo sapiens KIAA0402 mRNA, partial cds" (GenBank Accession No. AB007862)

· "v-kit Hardy-Zuckerman 4 feline sarcoma viral oncogene homolog (GenBank Accession No. X06182)"

· "catenin (cadherin-associated protein), beta 1 (88kD)" (GenBank Accession No. X87838)

· 配列表の配列番号 1 1 5 に記載された塩基配列を有する遺伝子

· 配列表の配列番号 4 3 に記載された塩基配列を有する遺伝子

· "Homo sapiens nibrin (NBS) mRNA, complete cds" (GenBank Accession No. AF051334)

· 配列表の配列番号 7 1 に記載された塩基配列を有する遺伝子

· "Homo sapiens ubiquitin hydrolyzing enzyme I (UBH1) mRNA, partial cds" (GenBank Accession No. AF022789 52)

· 配列表の配列番号 8 5 に記載された塩基配列を有する遺伝子

第4群 (24種)

- "integrin, alpha 7" (GenBank Accession No. AF032108)

· 配列表の配列番号 1 6 6 に記載された塩基配列を有する遺伝子

· 配列表の配列番号 3 6 に記載された塩基配列を有する遺伝子

· 配列表の配列番号 5 2 に記載された塩基配列を有する遺伝子

· "leukocyte antigen related protein (GenBank Accession No. X54890)"

· "CASP2 and RIPK1 domain containing adaptor with death domain (GenBank Accession No. U84388)"

· "superoxide dismutase 1, soluble (amyotrophic lateral sclerosis 1 (adult))" (GenBank Accession No. X02317)

· "wingless-type MMTV integration site family member 2 (GenBank Accession No. X07876)"

· "37 kDa leucine-rich repeat (LRR) protein (GenBank Accession No. U32907)"

· 配列表の配列番号 4 5 に記載された塩基配列を有する遺伝子

· 配列表の配列番号 1 6 7 に記載された塩基配列を有する遺伝子

· "v-Ki-ras2 Kirsten rat sarcoma 2 viral oncogene homolog (GenBank Accession No. M54968)"

· "v-jun avian sarcoma virus 17 oncogene homolog (GenBank Accession No. J04111)"

· "deoxyribonuclease I-like 1 (GenBank Accession No. X90392)"

· 配列表の配列番号 1 2 9 に記載された塩基配列を有する遺伝子

· 配列表の配列番号 2 0 に記載された塩基配列を有する

遺伝子

· "cell adhesion molecule with homology to L1CAM (L1 homologue of L1) (GenBank Accession No. AF002246)"

· "catenin (cadherin-associated protein), beta 1 (88kD)" (GenBank Accession No. X87838)

· 配列表の配列番号 1 1 5 に記載された塩基配列を有する遺伝子

· 配列表の配列番号 4 3 に記載された塩基配列を有する遺伝子

· "Homo sapiens nibrin (NBS) mRNA, complete cds" (GenBank Accession No. AF051334)

· 配列表の配列番号 7 1 に記載された塩基配列を有する遺伝子

· "Homo sapiens ubiquitin hydrolyzing enzyme I (UBH1) mRNA, partial cds" (GenBank Accession No. AF022789 52)

· 配列表の配列番号 8 5 に記載された塩基配列を有する遺伝子

第5群 (24種)

· 配列表の配列番号 1 6 6 に記載された塩基配列を有する遺伝子

· "Human beta adaptin mRNA, complete cds" (GenBank Accession No. M34175)"

· 配列表の配列番号 1 6 9 に記載された塩基配列を有する遺伝子

· 配列表の配列番号 1 2 3 に記載された塩基配列を有する遺伝子 (ID No. 123)

· "forkhead box C1 (GenBank Accession No. AF048693)"

· "insulin-like growth factor binding protein 6 (GenBank Accession No. M62402)"

· "cell division cycle 10 (homologous to CDC10 of S. cerevisiae) (GenBank Accession No. S72008)"

· 配列表の配列番号 4 5 に記載された塩基配列を有する遺伝子

· 配列表の配列番号 5 4 に記載された塩基配列を有する遺伝子

· 配列表の配列番号 8 6 に記載された塩基配列を有する遺伝子

· 配列表の配列番号 1 6 7 に記載された塩基配列を有する遺伝子

· "integrin, alpha 2 (CD49B, alpha 2 subunit of VLA-2 receptor)" (GenBank Accession No. X17033)"

· "tyrosine kinase 2 (GenBank Accession No. X54637)"

· "Homo sapiens unknown mRNA, complete cds" (GenBank Accession No. AF047439)"

· "465 (ID No. 155)"

· "Homo sapiens CGI-65 protein mRNA, complete cds" (GenBank Accession No. AF151823)"

· 配列表の配列番号 1 7 0 に記載された塩基配列を有する

- posis type 1)" (GenBank Accession No. NM_000251)
 - ubiquitin-conjugating enzyme E2A (RAD6 homolog) (GenBank Accession No. M74524)
 - "collagen, type VI, alpha 3" (GenBank Accession No. X52022)
 - low density lipoprotein-related protein-associated protein 1 (alpha-2-macroglobulin receptor-associated protein 1) (GenBank Accession No. M63959)
 - hidogen (enactin) (GenBank Accession No. M30269)
 - "cadherin 6, K-cadherin (fetal kidney)" (GenBank Accession No. D31784)
 - "collagen, type III, alpha 1 (Ehlers-Danlos syndrome type IV, autosomal dominant)" (GenBank Accession No. X14420)

第8群 (20種)

- insulin-like growth factor binding protein 3 (GenBank Accession No. M35878)
- "Human transactivator protein (CREB) mRNA, complete cds" (GenBank Accession No. M27691)
- 配列表の配列番号6に記載された塩基配列を有する遺伝子
- protein tyrosine kinase 2 beta (GenBank Accession No. U43522)
- tissue inhibitor of metalloproteinase 2 (GenBank Accession No. U44385)
- Rho6 protein (GenBank Accession No. U69563)
- "integrin, alpha 3 (antigen CD49C, alpha 3 subunit of VLA-3 receptor)" (GenBank Accession No. M59911)
- heat shock protein 75 (GenBank Accession No. AF043254)
- Homo sapiens mRNA from HIV-associated non-Hodgkin's lymphoma (clone h12-253) (GenBank Accession No. Y17173)
- "Human cyclin-dependent kinase inhibitor p27kip1 mRNA, complete cds" (GenBank Accession No. U10906)
- bone morphogenetic protein 4 (GenBank Accession No. U43842)
- tyrosine-protein kinase receptor tyro3 precursor; tyrosine-protein kinase rse; tyrosine-protein kinase

- * inase sky; tyrosine-protein kinase dtk (GenBank Accession No. U18934)
- metastasis associated 1 (GenBank Accession No. U35113)
- Human mRNA for SB classII histocompatibility antigen alpha-chain (GenBank Accession No. X03100)
- "matrix metalloproteinase 7 (matrilysin, uterine)" (GenBank Accession No. Z11887)
- "mutS (E. coli) homolog 2 (colon cancer, nonpolyposis type 1)" (GenBank Accession No. NM_000251)
- 10 · 配列表の配列番号139に記載された塩基配列を有する遺伝子
- 配列表の配列番号79に記載された塩基配列を有する遺伝子
- AG26 (WO 98/37187号国際公開パンフレット)
- "cyclin-dependent kinase 5, regulatory subunit 1 (p35)" (GenBank Accession No. X80343)
- 20 [実施例3] DNAアレイの作製
- 实施例2に示された第1群～第8群のそれぞれの遺伝子の組み合わせについて、それぞれの遺伝子断片を調製し、これをスライドグラスに固定化したDNAアレイを作製した。DNAアレイは第1群～第8群のそれぞれについて別個に作製した。また、DNA断片の調製～固定化の操作は実施例1-(2)に記載の方法で実施した。
- [0112] 上記のDNAアレイを使用して、胃癌試料における遺伝子発現プロファイルデータの測定を行い、さらに各試料について存命月数の推定を行うことが可能であった。
- 30 [0113]
- 【発明の効果】この発明の装置及び方法を使用することにより、遺伝子発現プロファイルデータを得た多数の遺伝子より、目的とする評価指標データを得るために最適な遺伝子発現プロファイルデータの組み合わせを生成することが可能になる。被検試料中の選択された遺伝子の遺伝子発現プロファイルデータを測定することにより、被検体の評価指標データ、例えば、余命、健康状態、疾患の分類、薬効の評価等を得ることができ、この発明は生化学、医学の分野において極めて有用である。
- 40 [0114]
- 【配列表】

SEQUENCE LISTING

<110> TAKARA SHUZO CO., LTD.
 <110> MITSUBISHI SPACE SOFTWARE CO., LTD.
 <120> Method for analysis of gene expression profile and a system therefor
 or
 <130> 01-0078
 <160> 171
 <210> 1

39

40

<211> 540

<212> DNA

<213> Homo sapiens

<400> 1

gatcaagtcc atacccagca gctatcaagg atttcaagg ctctccctac atctccctcc 60
 acctactcct gtttagaaggc ttgaaattgt gctgtaaatg tacgtgcata gttatagctc 120
 caccattnag ttcatgtccat tccctcagcc caaattctc tttccctcat tccttgcctq 180
 ttgaaagttt tcttttctt caaaaccacg agagcgccta ttcccttctg aagttttcc 240
 tgatattccc aqgcaagatc actcaactcat tccttcaacc atgtacccgt tgggtgcact 300
 tggatctcg tttgggtgtt ttgtcagattt gcatttctgc agacttagtt gttgtcacat 360
 ctcttacttt cattagattt taagcacattt aaaggcaagg gctgtatccc catcatatct 420
 tctagaagta ctacctaaca caattgccaag tacagatgt atgctcaata agttgtcat 480
 taaattaaat ttcaaggaa gtaatgtcat aaacacccctt agttaaaqca ctgagacacg 540

[0115]

<210> 2

<211> 536

<212> DNA

<213> Homo sapiens

<400> 2

gatcaagtcc aacaaaaca tgaaaagaaa gacacacatg gccaagtaca ttgtccttcc 60
 caqqaaggta agtaaagggtt gtttaacatc caaaaatqca atqaatgttttcaacat 120
 ttcttaacta aaaagaaaaa ccacacaaca atctctaaag tggcagaaaa aqcatgttgc 180
 aaaatccacg atgtataactt gattaacact ctcacagctt aatagaataa aacttccatca 240
 agctgataaa qgacatctac aaaaacccta caattnacat gacacttaat aatqaaggc 300
 tgaatqattt tcttctaaaga tcaagaacaa gccaaggatg tggatctactt actactaattc 360
 aatqttattt tggatgttctt aqccaggqca gtaaggcaag aaaaqcaac aagtataqta 420
 ataatactgt cttgatqggc agatgacatg tttgtctatg aqaaaattc aatqaatct 480
 ttaaaagaaac tcttqaaact aacaaggtagg ttttgcaaaat cttatccatca tcaacg 536

[0116]

<210> 3

<211> 511

<212> DNA

<213> Homo sapiens

<400> 3

ggattttta gtagaagacgg ggtttccacgg tgtagcgcgg gatgggtctca atcttctgac 60
 ctcatgtatcc gcccacctca qcctcccaaa qcactggat tacaggcatg aqccactqcg 120
 tctggccgaa aaataqctgt ttttttaaaa qcccaacatt atqctgaaac atggctaaat 180
 ggttaactta cactttaaat tattcactgg tgtagcaataa tacaacccaa aattttgtta 240
 tctagctgtt agtccctctcc tacttctt atqctgtct gctttggatc ggtccagtct 300
 atttggqaa atcagaccac aatqcttca aattactatt acaacacaa atcagatcac 360
 aqaaacagcca aaccccatc atatqgtcaaa aaagttacat tttttctca taaaacagaa 420
 aataqctcaat cttggtaaaa qaaaacatca gttcttggatc gatgaggatt cctctccaca 480
 tggatggatggaa taqcacacca qgcaaatcaa q 511

[0117]

<210> 4

<211> 404

<212> DNA

<213> Homo sapiens

<400> 4

cgagacagaa tctcaactctg tcacccaggc tggatgtccatg tggatgtcatc tcgtctcact 60

41

gctacctcca ctcctgggt tcaagcaaat ctcatgcctc agcctcctga gaaqctggga 120
 ttgttagat aqggttcac catgttggcc aggtcgctc cgtactcctg agtcaggta 180
 atccgcctgc ctcggcctcc caaagcqctg qqattacagg cgtqagccac cgcacccagc 240
 tqcaagtttq attttaaca tqctaaattt qagggtctt taggacatcc atgtqaggc 300
 atctqggagg caqctgacaa tctqgaggta tggagatq tcaqggtqa aqacatcaac 360
 agagggtatq qctgaaatct tgggaggagg taaaqgcaat caag 404

[0 1 1 8]

<210> 5
 <211> 157
 <212> DNA
 <213> Homo sapiens
 <400> 5
 qtaaatqata qcatqatatac aatqgcaaat qcgqggaaq tagacaaatc gattaccc 60
 tcggcctttq ttqcttctgt ggcaagtcc cccactcatc tctttacatt ttagtcaccc 120
 aaaaagaata atcataactg aaatcctqgc aatcaag 157

[0 1 1 9]

<210> 6
 <211> 364
 <212> DNA
 <213> Homo sapiens
 <400> 6
 cgagaaggag ttcactttq ttgcccaggc tggagtqcaq cgggtcaatt tcagtcact 60
 qcaacctcca cctcccaqgtt acatqcaatt cttgtgcctc agcctcctga qtaqgtggga 120
 ttacaggtag ccggccaccac qccctqctta tttttgtatt tttagaaagag atagggtttt 180
 qccatqtttag ccaggctggt ctcgaactcc tgacccagg tgatctgcct qccctggccc 240
 cccaaagtgc tgggattata qgagtaaqcc accgcqcccg qcccaataaa qctttttta 300
 agtgatcatq tctqaaactt acagtaatq qgctttacca cattccatq tgaccqacg 360
 ttq 364

[0 1 2 0]

<210> 7
 <211> 625
 <212> DNA
 <213> Homo sapiens
 <400> 7
 ctccacccccc caqccctcaaq caatccattt acctcaggccc tccqagtaqg tggggctaca 60
 ggtgtgttgc acaccaggcac qcccggttaa tttttgtacc ccctgttagq acgggggtttt 120
 qcttttqcca tgggtcccaq gctggcttag agctcctgg ctccaggcaat ccacctqct 180
 cgggctccca aqgtqctqgg attacaggca tqagaaacca cactcaggct agtaccccat 240
 tccttattat tggtaataa tatttcatttq tatttttttttccatqatqcc 300
 ttcatcgtt gatggacatt tgggtgttt ccactctgaa ctattatqaa taatqtcct 360
 atgaacattt atgtacaggat ctttqtaa catatqtttc catttctctt qggcatatat 420
 cttaggatgg aattccctggg tctttqctta accttttqag aacctqccaa gatgtttcc 480
 aaagcqctq taccatqtttta cattccacc accaagttat aqatqttca qttactccac 540
 atatattatq tctcttccc ttttttctt ttttqataca qagtcttqct ctgtcacc 600
 qaatqaaagac aqgtqggcaaa tcaag 625

[0 1 2 1]

<210> 8
 <211> 457
 <212> DNA
 <213> Homo sapiens

43

44

<400> 8

cgagatgaaq tctcaactctg ttgcccaggc tggagtgcag tggcaactatc tcaagctcaact 60
 qaaacacctca cctcccaagg tcaagcaatt ctccctccctt aqccctccaa gtagctggaa 120
 ttacaagggtg gcgccaccat acacggctaa tttttatatt ttgagtagag acagggttcc 180
 qccatgttgg ccagggccgt ttcaaaactcc tgacctcaag tgatctgccc tttcagcc 240
 ctcaaaagtac tgggattata ggcgtgaccc accacacccca qcctttttt cttacttag 300
 aqctaagacc aaacaatgca agggaaacaaa gtttaatatt aatgagcacc gcccacatgg 360
 catcaaaccc ttagtgcacgt tccatqaaqc tatgattaca attcctattt tataaatgaa 420
 gaaatgaaag ctcaagggtcc ctgttcacg gtcacct 457

[0122]

10

<210> 9

<211> 457

<212> DNA

<213> Homo sapiens

<400> 9

aggtaaccgt acagttatca gatgataaaat tctaattatcg qgacattcgc gtgggtcg 60
 tggtaattgaa gacagtgcac cctttgtac acgaaatgtgt gcaatgtgg acaggagtg 120
 aqggacagaa aqcgaaqgaaqg tggtaacgtg ctcttcttc actgtaccc tcataatccgc 180
 agccggatc agtcaagatc tcatttcgtg tgcttgggtg ctttggggc ttggaaagcc 240
 aqtcacggcc ttgtgtatcca ctcggattag tggtgtctgg ggttctact ttctgactcc 300
 tggaggacga aacgcccactt gtctgagggg gagggtggagt gattcttqag aatgaaaaca 360
 tgatgggtt agtgaagagg tggagggggg aggcaagggg gatgtggaaa aqaggttagaa 420
 gcaaggaaatg aqcatgaaat aaqgatgaaat ttggggc 457

[0123]

<210> 10

<211> 402

<212> DNA

<213> Homo sapiens

<400> 10

cggcaagtca aaccagagaa aatggtgaaa cttttaaatgtg acagtcgtact cgagacagac 60
 gggccctatacg ctgtggaccc aggttaagggtt agatgatcaa tggatgtgg caccactctg 120
 caacgggtgcg acttctccag aagtttgcg aqggggaaag agaaatgttatt aqagtatggcc 180
 acatgtgtgt ctgttggggaa aacqcccgacg accaacgcgg aqatggcggt aaacaaccc 240
 aqaaagagact gggctgcctc ctgttgcattc attccatca ctttttctt tcttgggttt 300
 atcccttaacc acgaaagagca gtttcaccc ttaatgttgc ctcaacatcc cttttgtata 360
 ataaataactt gcttagcgtc tactttact gcccacgtt tg 402

[0124]

<210> 11

<211> 501

<212> DNA

<213> Homo sapiens

<400> 11

cgttagagacq gggtttaccat atgttagcca gqatggctt qatctccgt tctcatgtac 60
 caccggccctc ggcccccgg aqgtgtgggaa ttacaggcgt gggccaccgc gcttggccgt 120
 gcttggccaa gttttctatt ttttagtggg acagagtttc accatgttgg ccaggcttgg 180
 ctgcgttccatc tgacccatcc tcccttggccct cccaaatgtgc tggatgtca 240
 ggtgtqagcc accatqccca gcccgttatct atatatttca aatqatqaaat ttaaqgaaat 300
 gaggaaattgtg gggccggactt aqagagcaat ggttgcgtat aqgtggatggc tctggctgg 360
 tggatgtatct tcatatgtaaat cccatgaaaa tggatgtatcc tcccaatgtac aacctacagg 420
 tggatgtatct tcatatgtaaat cccatgaaaa tggatgtatcc tcccaatgtac aacctacagg 480

45

caagccccat gacggtcacc t

[0125]

<210> 12

<211> 502

<212> DNA

<213> Homo sapiens

<400> 12

agqaacactq aatqaagttt attqaagaga aacacqtaact gaaaacqttq aaqaqaqqqq 60
 aatcacaaca caqqaqaaaa ctgcaaaacgg ccaccacagg acacagggtcc cacacagggtc 120
 ccggggccccc tcctcctggg gttcctggca aqacqccctca tctgctccca ctccctgctgc 180
 aqccccccgt qccaactccc tccttcgtqa qggccccccgg gtctgcctct tgccaqgaaq 240
 gttagactctq qcaqgttcca ctggctttag qgaqccccct cccaaqccca tctgctggga 300
 caqggggccctq ccctatqgccc cgcacacggac aqtcacacgga cactqaaaaa cacagcaggg 360
 aacaccctac cagggtcaagg aaactqagtc atqagtaaca acagggggtgg aacaccaaag 420
 cacaagctc ctcqggqagcc tcaqggcaggq caqccccqct ccqtcctqgg tccagttctt 480
 cctcccacag tcacqgtcac ct 502

[0126]

<210> 13

<211> 393

<212> DNA

<213> Homo sapiens

<400> 13

caaaccgtcq qtaqqaatqc tqatcatqat aqttttqggtt ttctacaqat tctqttccgg 60
 tgcccttcct atccaggcac cacctqagaa aqttgtcatt tgaggcgtca ctggaaagtt 120
 acatctgtqa agtttctgtc attcgtccag atctgtgtgt tgtagcatgtq ctgaggaaagc 180
 acgtgtctggg ctgtgtcctca gacagtgtcat caccggggcac ccaqaggqttt qccctggctat 240
 tcctgttctq tggtgtgtgg aqgtgtgggg aggaacagat qcaqatcaac ctgtggctgt 300
 ttccctgtct aqgttctcac aqgtatctcc tqacacqggtt acttaacaat qgctctgtq 360
 gaaatttctta taaaataat qtcacaaaatq qcg 393

[0127]

30

<210> 14

<211> 381

<212> DNA

<213> Homo sapiens

<400> 14

gatcatqgtc acaaacacgg gtaaattctq tggattaaat aaqgatatct aqagatctac 60
 aaacacaaat ctaaagaatc cttgatttca ctggaaatca aqatgttga tgaaggtaat 120
 catqgatctat tggttttac tcataaaaact qgqggtgaaa qacaaggatt aacgtttctta 180
 qcgatqgttaa aqgtgtaaaa aqgttttgc tgqgqgqgtq tggtgtataa atttttacaa 240
 ccttctcaga aaacaatatq gaaacgttcca qcaacaqcca qcaqgaaaag cctgtttqag 300
 qccctgtgtq tggtataqggg tggctgtctq aacacagggg aatcatgtq atccatgtca 360
 qctgqagtca ttttggtaat c 381

[0128]

<210> 15

<211> 266

<212> DNA

<213> Homo sapiens

<400> 15

cgggggggaaag tggataaat tccaaaggaa atttaaaaaa caaatacatq aaataaaaaa 60
 ttgtgtcagat agtacagtac aqggacaaag atcccatagt qggagtcaaa ttatctgac 120

46

501

47

48

tcaaatctgg ttctgccact tactgacttt ggatgagctt cttAACCTCC ctggcctca 180
 attttaagaa ttaagtqaaa caacccaaga aaaatttagc acagtactga gagacagagg 240
 cacacacaca agcacagacc atgatc 266

[0 1 2 9]

<210> 16
 <211> 523
 <212> DNA
 <213> Homo sapiens
 <400> 16
 gatccatqgt cttttcttaq aqgatctttt qaacactaqq aqggacattq agatccccgt 60
 taacccctca qacaactaca cataaaacat qctagataqg tgataactca qtttqtaaq 120
 cactqagtgt tctctctgtc ccacacccag ttctcttagac tctaggqata cacagataaa 180
 taagtataqg tcttqttctt qtagtqttatc aqgtqtttag qaaatatttq aaaqqacatt 240
 ttataaacct ctagtttagc atttgttacc caaatgttatt ctcttttct qcacccgt 300
 qcaattqaaa atqgqtaatg cccatttct qaqaaaqgqtc cttqatattc tcagaaaaat 360
 tattgttqaa ttctcttcc tgggttaaqg aataccqgtc cattaaacat ttcttcataa 420
 aattcctqgc tagatqgatc aattttataga cctctcttatt ccatgagaat qaacagacat 480
 tqagacataa ttttatacat cagtataqaa taggagtcac ccq 523

[0 1 3 0]

<210> 17
 <211> 611
 <212> DNA
 <213> Homo sapiens
 <220>
 <221> variation
 <222> 165, 486, 503, 504, 512, 584, 599, 609
 <223> n is a or c or g or t.
 <400> 17
 ttttggctcc gtgggacgtt gtaatgtqca cagacatttc caagggaaattt ctaaacagtc 60
 accctccct tttgcattcc cccaaatctt aqgtgtatc ataaaacccct gtgtacatat 120
 tgggtqgtq ataaaaaggaa attggtaaa cagtagactt qggtnatqga actttctgtq 180
 qccacccatq aagacaagt taacaaactq tcatggaggg ctgttqttgc ccagccaggq 240
 ccgctqcat tttgacaacat ttccaccctg gccactqgc acatttcattq qaggcatgt 300
 cttttcaatq atactttttt qataqttttt atataacaaa atccttattc tatttataac 360
 ttaagatqat aaggcactat aaattaatqg cctaaaataa tatatttqtc tggtatctt 420
 tgcttatttctt acttcacttt aatttttagc tgtaaaattq gtaaatggat tcttacqact 480
 atctcncccc cccccccccc ccnnqgtttt taaaatttaa ttaaatqgca gaatttcctt 540
 tttttaggaa caatgtgtat tttcaattqgc acctatttaa aaantaaact caaaatttng 600
 catgccccgna a 611

[0 1 3 1]

40

<210> 18
 <211> 480
 <212> DNA
 <213> Homo sapiens
 <400> 18
 gttttcgcaq cacagtctta tccagttcca ttattggggta qcaaaqatq agtaaggaat 60
 qaaaacagga caaaaacctcc qctcttqatq atcttcattt aqccaaqaaq aqtaaqaaa 120
 aqagttaaca tcaatcatga qgtcctggaa qqaqataaq ctctqatqtc aaatqacct 180
 qggatccacc agtaatqatq gaaatttaga aqtaactta aacctctcaq gttcagttc 240
 ctcatctata qgaqgqaaat acaaattcatt acaaataacta qcaaactqaa tccagcqca 300

49

caccataaaa ttaattcacc atgatcaat gggctttatt cctgggggtc aaatttgtt 360
caacatactc aaatcaqtaa atgtgattca ccacataaagc aqattaaaaa acaaaaacaga 420
catgatcatc tcaatagacq tagaataaagc ttttgataaa atttaaacatc atttcatgac 480

[0132]

<210> 19
<211> 333
<212> DNA
<213> Homo sapiens
<400> 19
cgagtaaaat agaaaattctt taatatacag tqcagtttt ataaaagatgc tctgtgtgtt 60
atgctgaaacg cacaacatc ctttagaaaaat agtaatgtaa aaattcaaaa qgqttccctt 120
attgcagata ttcttgqaga tqtttaattta tgcatctctt cqaaaggcga cagccatcca 180
accatgttct ctttcttta caacagcttc aaacatttctt cttctqagat qcatatttgg 240
qaaatgttgg aataaaattat atccagtctt tcataatttgc atctatgcaa atctgtqaa 300
aaqcttaaot ttatttattta caqctqcgaa aac 333

{0133}

<210> 20
<211> 243
<212> DNA
<213> Homo sapiens
<400> 20
cgtaa~~g~~act ggagg~~t~~gttt ~~g~~ggaa~~g~~t~~g~~cc tc~~a~~gac~~a~~c~~g~~c t~~g~~tt~~c~~t~~c~~ta~~g~~ c~~c~~at~~g~~g~~a~~ta 60
ccat~~t~~c~~c~~ta cctact~~t~~ccc g~~t~~cacatt~~c~~ c~~t~~gt~~g~~caa~~a~~ac cccat~~t~~gtt ctctaa~~a~~ct 120
cg~~c~~tc~~a~~tg~~c~~a gt~~c~~tac~~cc~~ct c~~t~~ccc~~c~~agt~~t~~ cccat~~g~~t~~c~~gt t~~g~~ttt~~c~~t~~c~~t tcc~~c~~ta~~g~~c 180
cttaaccatt t~~g~~g~~c~~t~~c~~ct~~g~~a att~~g~~taatt~~c~~ t~~c~~t~~g~~ccc~~c~~tg t~~g~~t~~c~~t~~c~~t~~t~~ g~~t~~c~~c~~t~~g~~caa 240
aac 243

[0 1 3 4]

<210> 21
<211> 857
<212> DNA
<213> Homo sapiens
<400> 21
gggtgcagac agggtttac catgttgccc aggtggcct tgaactcctg aqctcaaatq 60
agtgttagga ttataggca tgaqacactg tgcacgcca tcattacata atattacatc 120
agtcaattat qtaactaaaa ccaaaaqat tttaqatct ctttcaccc ccaqgcaaac 180
caagaacact qataattgag tgaaggaggt catttttaag qatttaggct ccatagtgaa 240
ttcacaataa ttccaaaaga catcaccctc agggcaattt tccaaaagta ctatactatg 300
ctgccttca aatagcgtaa tattaaaccq atgattccag tatgaatcct ctagaactaq 360
aaccacattn qcttaataa aaacaattt qacaaagtcc ttggttgtcc aacttgttcc 420
ctcaagtcgt tggatcttt actttcttc tagtctttt ctaattcagt tqaqtaaaaa 480
tcagttgaa aaattaacca gtcaggtaact aactataaag ttctgttaat acttcctatc 540
atcattcaact gtttaatca caccacactg ggaaatcaac acttqtaaca qacacgctat 600
tccaaaaaqc aaqattcatt tggctcctt qaaaaqaact caaaatgaaq agaaaqccatt 660
ttcattaact cctaattac acataagttt atcttttaca gtatattaat gctatTTG 720
atttttqagt aqggatttgg qgatgtcgtc aaataqctcc aaactcaaag tgcacccatq 780
cagcagcaga ggcattgagcc tttacacaag tqaqtcagq qgaatccatq ctcttqatq 840
tctqaaccq acqtttq 857

〔0135〕

<210> 22
<211> 786

50

ggtt

51

52

<212> DNA

<213> Homo sapiens

<400> 22

ggggtcgaaa atatgtctta ttcacttttc aagacttttt tcctgtgtttt taaaaaaatq 60
 aattgaacgg aattgattgg catcacctgatgacta ggacattqaa gatccattt 120
 cgtcactcat ctttgtqaga aagacactaa tacgataatq atataaccac cttcattttt 180
 gtatgcttag cagatactag accatgcatt cagcaccatc attagtatta tctcatttat 240
 gatTTTaaac aactggaaqqq ttttaggtact gttatcatcc ctacttaaca cagattgaaa 300
 aacttccata aqagcccaaq ctggtaaqct gtggqactqaa gctagatttc aaattcagat 360
 ctgtctgact ccaagtcagt gttctgtcct ataatgttct acagtactc tcaatttagt 420
 ggtqagacgt tagaaqctac aggtaaaaat ggctagtcctt caaaagttcta ccaactcaat 480
 aagtgcacat aatttagtgg gatattgtta gqaataaaag gatgttcaca gcttggaaat 540
 ttcaatatct aaaactccccc taqqcacctg acctttcctg gctgttaggca ttactccaa 600
 gaccagtaat ccaggaaccc agataatcaa tcacccacag cacaqcttcc ttgttgcatt 660
 caccatggct tttcaqagca tctgtcacca tgctttqtcata cacaataga gatgtcagtc 720
 ttcaactccg tgaaagttaa tttctctgct ctttgctaaa aacaggqctc tttctactgc 780
 gaaaac 786

[0136]

<210> 23

<211> 782

<212> DNA

<213> Homo sapiens

<400> 23

cgtagaaaatq gggcctcaact aggttgcaca qgctggctc caactcctgg acttaagtq 60
 tcctcctgccc ttagcttcctt aaaaatgttgg gattacagt gtggatgtact atgcccagcc 120
 taaaaatata aqccacaga ttatccacag qctaatttagt qaaaacagat tgggttat 180
 tggggactt agtataqaaac aqggcttqat tcataaaccac qggaaaatq agqactqat 240
 tggtaataaa atgattqact atcaggtaaa tagaaatcaga agtgtatgtt gaccctctt 300
 tctataccaa aagtatacta atgatttcaa aagttaatgt aataaaacaaa taaaggctaa 360
 gaaaaataat aqgtttat acatgtttt ttatgttcaq caattgtaaq gacctgtq 420
 acaaataaaa qcaatggaaq qcaaaaatta acaactttqa cacagcattt ttttatctqa 480
 atatcataaa tcaqaaacaa aattaaaaaga caaaactgtq gggaaaccqaa taacaaatata 540
 gqcaaaqaat taaaaccctg aagaagtcat gtagatcaa gggtaacattt aaaataaaag 600
 ctcaataggg aaattqacaa aqggggqgaa cataatctat ctaaaaqaa atacagataa 660
 ccagcaaaca aatatactca gcttcacaga gaaatgcaaa ttgggtgttt tttctatc 720
 ctttatttaa aqattttaa aqgtqactac tatacccaq tgggttaaggq gctgqcaaa 780
 ac 782

[0137]

<210> 24

<211> 539

<212> DNA

<213> Homo sapiens

<400> 24

cgtgacggag ttcgcactg ttggccctggc tggagtgcaaa tggcgtggc tcggctca 60
 gcaacccctggc cctcccatgt tcacgcccatt ctccctggc agccttcgtq gtagctggaa 120
 ctacggggcgc ccaccaccaat gcttggctaa ttttgggtt ttttagtqaa gatgggggttt 180
 caccatqttt qccaqgatqg tcttgatctc ctggcccttq gatccacccq cctcggcc 240
 ccaaaqgtct qggattacaq gtgttaqctaa ccacqcccaq ccacaaaaaq cctgttttat 300
 aataagggtgt tttttagtqaa ttttgggtt ttttagtqaa gatgggggttt ttttagtqaa 360
 gagacaggtt cacaccattq taaaqgtcaaa aagtcttaq ttggaccacc caaaqgtcaag 420

53

54

gactgtctat atagcctatt gttcctgttt ctctgaagaa ccctaactag cacagcgta 480
tccactagat tggatgaaag ataatgtgtc ttggacaaag gtggttgttt gcgaaaacc 539

[0138]

<210> 25

<211> 534

<212> DNA

<213> Homo sapiens

<400> 25

gttttcgcag taacatacca ttttagttca ttttacaaat gagaaaaatgg gttatacagg 60
aagtacataa gttatgcaga gtcacttacg agtggtagta ccgtgtcaag ctctatggct 120
ctagcatcg tgcctaaact cctgtgttac atggcctatc aaataatqag gaaaggaaat 180
ttcaaaaatt gtttgcctt taactgggtg gaaaaaqaat cacaatqaat cacataaatt 240
ttttctattt gttgttagat atccttctgt atctccgtgt acaatctctg ccccaaaagt 300
agtggggattt caattcagaa aagaaaagata attaaaaaaa agataaaaaca attaccctta 360
gttacacaat tcattgttaa catgttataa ttgttttaag gaacacattt atggcggac 420
atggtagcgc atgcctgtaa tcctggcgtt ttggagggcgtt ctaagggtggg aggttactt 480
gaggcagatg tttgtgacca gcctaggaaa tggatgtgaga gtgtatctt tacg 534

[0139]

<210> 26

<211> 322

<212> DNA

<213> Homo sapiens

<400> 26

gttttcgcag cagctgttct gtgttaggtgt gagaatgtgt ctttacagggt gggatttctg 60
tttggatgtg tggaaagga gaaaggcatat tgataaagcgt acaaacaqac atatctatqc 120
aaatacqca ctcgcagaaat aqgcqaggag agctagttt ttgcataaaa tcaataccqa 180
agacttggac atccagtcaa gaaaggqag qccaggcagc atggttcaca cttgtatcc 240
cagcacttqg qgaggctqgg qcaagggtggat tgcctqagct caggattcca agaccagaac 300
attgttaggac cctaaactcta cg 322

[0140]

30

<210> 27

<211> 206

<212> DNA

<213> Homo sapiens

<400> 27

cgcattcaca ttcaaggctat tcttaaagcgt ttttctqcca tggatcacatt gtgtatqaga 60
acatgtatgtt cactagtggg taactttctg tgcattgtcc ttactctca gggatgtcta 120
gtggattttac ctacccctgc ttttgcataa ccactgtaaa tctaataqgtg aaaaggcaaa 180
tgatgtctca gtatcactgc gaaaac 206

[0141]

40

<210> 28

<211> 479

<212> DNA

<213> Homo sapiens

<400> 28

cgaggtagag tcttactctg tcacccaaac tgggtgtgtgg tggtaagatc acggctact 60
gcagcttcaa tctcaqcgaa ctcaggtaat cttccaccc tggctccca agtgcgtggg 120
tcccaqggcc cacaccacca cacccaaactg attggatattt taggttagaga caqqatttca 180
ccatgttgcc caggctgttc ttgtactctt gggtcaagt gatccacctg ctttggcctc 240
ccaaagtgtctt aggattacac gataagcca ctgtgcctgg ctggatattt atctacttgt 300

55

56

aacctaacaa atta~~q~~cat~~q~~ a~~q~~g~~ttt~~c~~t~~g g~~o~~t~~q~~ct~~q~~ca q~~g~~aa~~ac~~c~~t~~g a~~q~~ttc~~c~~t~~g~~at 360
 a~~q~~agaaaaaa q~~g~~ac~~a~~g~~t~~tt~~a~~ t~~t~~t~~c~~cc~~c~~ag taat~~g~~at~~g~~gt a~~q~~cc~~a~~g~~a~~ct~~a~~ c~~c~~tt~~c~~at~~t~~tt 420
 t~~q~~cat~~q~~at~~t~~ c~~cc~~ca~~q~~t~~a~~cc~~a~~aaa q~~g~~aa~~at~~q~~g~~ga c~~a~~ct~~t~~q~~a~~ac~~c~~ t~~q~~cqaaa~~a~~ac 479

[0142]

<210> 29

<211> 877

<212> DNA

<213> Homo sapiens

<400> 29

c~~g~~ac~~a~~c~~t~~q~~g~~c a~~q~~c~~t~~t~~t~~c~~t~~ t~~t~~t~~c~~c~~t~~g a~~a~~ta~~q~~ag~~t~~cc caaaaat~~tt~~g c~~ca~~aa~~q~~gt~~tg~~c 60
 taat~~t~~ct~~c~~g~~t~~ t~~c~~c~~t~~t~~c~~aa a~~q~~qq~~a~~q~~c~~ac~~q~~ ct~~a~~ca~~t~~g~~a~~ a~~cc~~c~~t~~t~~c~~t~~c~~ c~~cc~~g~~c~~t~~tt~~tc 120
 a~~g~~t~~ttt~~t~~c~~c t~~q~~ata~~t~~t~~t~~cc ac~~g~~g~~a~~cat~~c~~ t~~t~~agg~~ca~~aa taat~~ttt~~t~~c~~c c~~cc~~t~~cc~~aa~~at~~ 180
 t~~c~~ag~~t~~t~~c~~g~~a~~ aa~~a~~cat~~c~~aa~~q~~ q~~c~~aga~~a~~cat~~g~~ c~~ag~~ct~~t~~t~~c~~g~~c~~ cat~~t~~tt~~ta~~g~~a~~c at~~cc~~at~~g~~g~~t~~a 240
 taat~~agg~~cc~~a~~ a~~g~~g~~at~~tt~~c~~c c~~ag~~ct~~c~~a~~g~~t~~a~~ caa~~g~~t~~g~~aa~~t~~ t~~t~~aaa~~a~~act~~q~~ ca~~a~~cat~~c~~aa~~a~~ 300
 at~~cat~~t~~t~~g~~a~~ g~~t~~t~~a~~c~~t~~t~~c~~g~~t~~ t~~q~~caa~~aa~~aca~~a~~ a~~ca~~aaa~~a~~ata~~a~~ a~~g~~t~~t~~t~~q~~g~~a~~ t~~g~~cc~~a~~tt~~t~~tt~~a~~ 360
 a~~c~~ag~~c~~t~~ta~~g~~t~~ a~~a~~c~~t~~at~~t~~g~~a~~ a~~ta~~t~~q~~t~~g~~tt~~t~~ t~~t~~c~~a~~c~~g~~t~~c~~t~~c~~ ac~~c~~t~~q~~t~~g~~ct~~a~~ g~~a~~t~~q~~g~~g~~at~~q~~g~~a~~ 420
 a~~q~~gg~~at~~t~~a~~ta ac~~g~~gg~~gg~~act~~t~~ t~~t~~agg~~g~~g~~a~~t~~c~~ c~~t~~g~~c~~c~~t~~t~~c~~t~~c~~ att~~a~~ca~~t~~g~~a~~ a~~at~~g~~c~~ag~~t~~a 480
 t~~q~~aa~~ta~~a~~a~~ a~~t~~ t~~c~~at~~t~~c~~t~~g~~a~~ c~~a~~taa~~q~~g~~a~~cc~~a~~ c~~c~~tt~~q~~g~~a~~act~~t~~ a~~a~~cc~~q~~gg~~g~~aa~~t~~ att~~ca~~t~~c~~act~~a~~ 540
 aa~~a~~at~~q~~t~~g~~t~~a~~ t~~c~~ag~~c~~acc~~t~~ t~~c~~t~~g~~c~~t~~g~~a~~ aa~~g~~gg~~ac~~aga~~a~~ a~~q~~ag~~t~~ta~~a~~ c~~t~~cc~~c~~t~~cc~~ag~~a~~ 600
 t~~q~~aaa~~aa~~q~~aa~~ t~~c~~aa~~ttt~~t~~c~~t~~c~~ t~~t~~tt~~c~~c~~t~~g~~a~~ g~~g~~at~~c~~q~~g~~g~~c~~c~~c~~ c~~a~~ct~~t~~t~~a~~ca~~a~~ t~~aa~~act~~t~~tt~~t~~ 660
 g~~t~~tt~~a~~at~~q~~g~~a~~ c~~c~~act~~t~~aca~~a~~ g~~g~~at~~q~~g~~a~~act~~c~~ c~~a~~ct~~t~~aa~~g~~g~~a~~ c~~c~~aca~~a~~act~~c~~ c~~q~~taa~~q~~g~~c~~at~~t~~ 720
 g~~t~~at~~gg~~tt~~t~~ at~~c~~t~~c~~t~~g~~t~~c~~ t~~c~~t~~at~~c~~t~~g~~c~~ t~~q~~c~~act~~t~~c~~ac~~t~~ tc~~ac~~cc~~c~~at~~c~~ t~~c~~cc~~q~~at~~c~~ag~~t~~ 780
 t~~t~~t~~g~~at~~t~~g~~a~~ t~~t~~ac~~at~~c~~c~~t~~c~~ t~~c~~q~~ttt~~q~~t~~ca~~a~~ g~~c~~act~~t~~ca~~a~~ c~~ca~~aa~~t~~tt~~q~~gg~~g~~ q~~g~~gat~~c~~ca~~a~~ 840
 t~~t~~g~~t~~ata~~at~~tt~~t~~ ca~~q~~ag~~t~~g~~a~~ t~~c~~ct~~agg~~g~~c~~t~~g~~ c~~g~~aaa~~a~~ac 877

[0143]

<210> 30

<211> 876

<212> DNA

<213> Homo sapiens

<220>

<221> variation

<222> 375

<223> n is a or c or q or t.

<400> 30

c~~g~~at~~q~~aaa~~a~~ a~~t~~q~~t~~q~~g~~aq~~q~~ q~~g~~ag~~t~~act~~t~~g~~a~~ g~~g~~ag~~gg~~g~~g~~g~~g~~ g~~t~~at~~ttt~~c~~g~~t~~t~~ t~~t~~tt~~ttt~~t~~g~~t~~a~~ 60
 t~~t~~t~~g~~tt~~t~~g~~a~~ t~~t~~g~~t~~ag~~q~~at~~t~~ g~~g~~gg~~g~~g~~g~~t~~c~~ t~~c~~a~~t~~g~~g~~ct~~t~~ a~~a~~ct~~c~~c~~t~~g~~g~~c 120
 ct~~c~~ta~~q~~g~~a~~at~~t~~ c~~c~~tt~~c~~g~~a~~ct~~t~~ t~~t~~gg~~c~~tt~~c~~ca~~a~~ a~~a~~gt~~q~~ct~~g~~g~~a~~ att~~a~~ca~~g~~g~~g~~g~~c~~ t~~t~~g~~c~~cc~~c~~ac~~ca~~ 180
 t~~q~~cca~~q~~g~~a~~cc~~a~~ a~~g~~a~~t~~tt~~t~~aa~~a~~ a~~aa~~at~~t~~g~~c~~cc~~a~~ a~~g~~gt~~q~~t~~g~~tt~~t~~ g~~t~~qc~~ac~~cc~~t~~ c~~q~~ag~~t~~cc~~c~~ag~~t~~ 240
 ct~~c~~tc~~t~~g~~c~~t~~t~~ q~~g~~ct~~g~~g~~g~~ca~~a~~ q~~g~~ag~~t~~g~~c~~tc~~c~~ at~~q~~g~~a~~cc~~a~~ g~~g~~at~~t~~ca~~q~~g~~a~~ t~~t~~ac~~q~~g~~g~~t~~g~~g~~a~~ 300
 ct~~t~~at~~q~~g~~c~~c~~t~~ at~~q~~t~~g~~t~~a~~ct~~c~~ c~~q~~g~~c~~ct~~g~~g~~g~~ca~~a~~ q~~g~~ct~~q~~g~~g~~g~~t~~g~~a~~ c~~c~~aca~~q~~t~~ta~~g~~a~~ acc~~c~~t~~g~~t~~c~~t~~c~~ 360
 t~~t~~aaa~~aa~~aa~~a~~g~~a~~ a~~a~~ta~~aa~~at~~q~~ct~~a~~ g~~g~~at~~q~~g~~a~~aa~~t~~at~~q~~g~~a~~ c~~t~~g~~g~~ac~~c~~ct~~t~~ t~~t~~ta~~ttt~~c~~t~~c~~t~~ 420
 ct~~c~~tc~~t~~g~~c~~t~~t~~g~~a~~ g~~t~~at~~ttt~~g~~c~~act~~t~~ a~~t~~ta~~ttt~~c~~t~~c~~t~~ t~~c~~aa~~q~~gg~~g~~at~~t~~ t~~t~~tt~~at~~g~~g~~t~~c~~t~~c~~ 480
 t~~t~~at~~ttt~~g~~g~~cc~~c~~ t~~t~~g~~g~~g~~g~~at~~t~~act~~t~~ a~~a~~gt~~q~~g~~g~~aa~~q~~g~~a~~at~~t~~ ta~~ac~~ata~~q~~at~~t~~g~~a~~ t~~t~~ct~~q~~tt~~g~~g~~g~~g~~a~~ 540
 at~~aa~~aa~~ac~~gg~~g~~c act~~t~~ac~~t~~g~~a~~ a~~ac~~at~~g~~gt~~t~~acc~~t~~ a~~aa~~ca~~q~~act~~g~~g~~a~~ t~~t~~g~~g~~ca~~q~~gt~~t~~g~~a~~ 600
 t~~t~~t~~g~~tt~~g~~g~~a~~at~~t~~ c~~at~~at~~t~~g~~a~~ a~~t~~t~~ttt~~q~~t~~g~~a~~ t~~t~~tt~~t~~ct~~act~~ t~~t~~g~~t~~act~~q~~g~~a~~ q~~g~~t~~q~~g~~c~~att~~t~~t~~c~~ 660
 t~~t~~g~~t~~g~~c~~cc~~c~~aaa~~a~~ a~~aa~~aa~~q~~t~~t~~t~~a~~ t~~t~~aaa~~aa~~aa~~a~~g~~a~~ a~~g~~aaa~~ac~~cc~~a~~g~~a~~ ac~~ac~~gt~~t~~ac~~c~~ a~~ca~~aa~~act~~tt~~q~~g~~a~~ 720
 q~~g~~q~~g~~ca~~q~~act~~t~~ g~~t~~tt~~q~~g~~a~~ata~~t~~ t~~c~~q~~at~~q~~g~~aaa~~a~~ t~~t~~g~~t~~ac~~t~~aca~~a~~ a~~aa~~g~~g~~act~~t~~ 780
 ag~~g~~acc~~at~~g~~a~~ g~~g~~ag~~t~~g~~g~~g~~a~~ a~~g~~ag~~g~~aa~~a~~ac~~t~~ c~~ac~~g~~c~~acc~~g~~g~~c~~ t~~t~~g~~g~~aa~~q~~g~~a~~ca~~a~~ cata~~aa~~at~~t~~g~~a~~ 840
 a~~aa~~g~~g~~aa~~a~~ac~~t~~g~~g~~g~~a~~at~~t~~ t~~t~~aa~~ac~~act~~t~~g~~c~~ q~~aa~~aa~~a~~ac 876

[0144]

50

57

<210> 31
 <211> 802
 <212> DNA
 <213> Homo sapiens
 <400> 31

gttttcgcaq cagtagaqaq ggacagcctg qggAACGCTG qggccctgtc gaatcccggg 60
 ctctccaaqc acqaaacaggq qcatcccgq tccacgcccq qgggaggaggq gtttcaccgt 120
 ctccaaqqac ttggctcqag gaaattaaac tctaattqagt aqctcqagaa atqttccggg 180
 atggtggatg agcgatcatc cttaaaqaa aatqctattc tgggagctcc aacctgcaat 240
 taacctacag aqgaaaccctt ttgagaggct qgtqcaqgc ttcggggagg cagattaaga 300
 acttagttgg cctgggtgaa qctctgtqag qaqcaaagca gccccttcca qgtqaactqc 360
 ttqactttac caccgtqaagg agtatttact qcaagaatta acaaagcagg tggtccgtat 420
 ccaagcagaa ccctgttatt tqaagqaccc qagaaatctt atcccgttta caaatgtcct 480
 qggagtgcac ccaaagtqaa acttgggtgg atqaaaccqca tcaaagactt ccagaagcca 540
 ctccaaqtgac cataaqataa atgtcaqaaat cqcatataaccc ccaggatcag agatqgagtc 600
 tcgctctgtc acccagcctg qaatqtagtq qcatcattgt aactcactqc aqccctgaaac 660
 tccttggctt aagtgtatcct cccacctcq cctcccaagt aqctggaaact ataggcaca 720
 qctaccatqc ctggctattc tttttcqttt attttttttta aactaaggaa tgaagaaaat 780
 aaaacattac atgactqgta cq 802

[0145]

20

<210> 32

<211> 783

<212> DNA

<213> Homo sapiens

<220>

<221> variation

<222> 663, 670, 679, 684, 688, 691, 692, 706, 714, 733,
 734, 775, 777

<223> n is a or c or g or t.

<400> 32

cgttagaaaca aqgtcttact atqttgcccq qgggtggctc caactcttgg qctcaagcaa 60
 tcctcctgccc ttggcccttc aaagtqccgg gattacaggq gtgagacact qcacccggcc 120
 ccttatttttq taactcttat caccattata acaacaccqaa tggtatcttt ttaatqtcgt 180
 caqtcgtccc tactaaactq caaqctgtq qgaggcaaga accatqctq ttcactqctg 240
 tatttctaqc accaagtqqa qcgtctqgag cacaqgtqag agtaaccatt ataaacqgtc 300
 ctcaccattq tcccaqccat ccagaataac caqctaactt cacqgatqta aqagqtaaac 360
 actttcqat cagttgggtt cacagatqgt caatattqtc agatqaaaag taaattcaac 420
 atcttcctgg tttccaaagg caqaatcaq aaaaacctqaa aaaaacaactq cqgaccctct 480
 qaatttaaac tgacaqctaa cttcagattc aaaatttggg aatcacqcc ataaaaccaa 540
 qcaaactagt agattqactc tggatqaaa caccatcatq qgtacqcc ctqccccc 600
 tctactcaat ctggacaaaa aqagggaaac aqgcaaatgt gtgtgtcttc tgcttgcct 660
 tantttcan cttccacnt ttantqgnac nnccgaaaac cctgggnqta aatngtaacc 720
 cqgtccaatt qgnngaccct caaggcatqc aqcttqgca ctggccqtcg taatngncat 780
 caa 783

[0146]

<210> 33

<211> 272

<212> DNA

<213> Homo sapiens

<400> 33

59

cgtaaatqcc tgatttata ctgttcatgt ctactttaa taaaacaaac aacatgttta 60
 atqacaatta cactccaaa gtqaggaa acggcaqqa cgtgccacct aaaaatcaqc 120
 cctgattctt ttctcaggc tcaaggctggc tttccctctt gataactgca atqctggct 180
 ttggaaqqqqa caaaaqgaga qcccacaggq tggaaaccg tctccata ttqctgtqga 240
 gtatgtgtgt cttaatttc ctggacttga tc 272

[0147]

<210> 34
 <211> 684
 <212> DNA
 <213> Homo sapiens
 <220>
 <221> variation
 <222> 663, 678
 <223> n is a or c or g or t.
 <400> 34

acaagggttt tcgcagatga gaagatgaga aaggctacta gcacgtgtgg cctgccacac 60
 agcaggcacc cagtqagggt ggactcaact tccatqaqc ataqaaggca tccgtctgtc 120
 ccaggccctt qgggtggcac ttgggccttt tggaaqtggc tggatgggttt agggaaatttg 180
 tcttcctcttq aatcaqctq acqgctcatq aatttqgtat atttaaatq ctctqgctgg 240
 qcattttggc tcatqctgt aatccccatq ctttgggagg tggaggaggq agqattqctt 300
 gaacccaggq gttcaagacc agcctgggca acagtqagat cgcgtctcta caaaaatttt 360
 tttaaaaaat tagtqggca tcaqgcaqgc qcgggtgtc acqccctqtaa tcccacgact 420
 ttgggaggcc qaggcaggcg gatcacaqg tcaaggatqatc gagaccatcc tggctagcac 480
 qgtaaaaca tqtctctact aaaaatacaa aaatttagctq qgcatqgtgg tgggacccctg 540
 tagtcccaqc tgctcgggag gctgaggcgg gagaatgggtg tgaacccggg aggcaqagct 600
 ggcaqgtqagc cgagatqgcq ccactgact ccagcctggg cgacagqgc agactccatt 660
 cgnaaaaagg atcaccqngt accg 684

[0148]

<210> 35
 <211> 536
 <212> DNA
 <213> Homo sapiens
 <400> 35
 gttttcqcaq ctactaaggta aaactqatct taagggtata qcaggcttatt tcaqcaqctq 60
 gcttgcaga taatccctgg qcaaggctttt qtgcccttaq tggatgggtttt ttcctttctt 120
 tcctccctca tctcatqact gtaatttat tggatgtac aagataattc caatttatat 180
 aatcaatttt cacaatatcc aatqcatcca aaattaattt ttctttatat ttttatgggt 240
 tacttqagtt ccatgtactt gttataatgtt attqaagttt ataattaatc tttcqatgt 300
 aaacttqqa taattqaat qttqaaqata tggatqactat attatqaaa tactcattgg 360
 qctatqcaqgtaatqatqctt cagacgtggat atqtaatgtt tggatggaaac aatatttcat 420
 gtcgggttcqg attcaqgttt ttatgggtt aacattqca tggaaatttt tatgtttaat 480
 tctgttagct qgtaacagaa cttctgggc aattttgttt caqtttqttq taactc 536

[0149]

<210> 36
 <211> 619
 <212> DNA
 <213> Homo sapiens
 <400> 36
 ggtaggaaat ggtqtccttac tggatggccag catqgtcacq aattccctggc ttcaaqgtat 60
 cctccctaccc caqccctccac cctqccctggg qtcctccctga qggctqggac tggatctgt 120

61

62

tgtgtgtact gccatattcc caaaacctaa acgatgcctt gcatggaaaca ggggatcagt 180
 ggtgattgt tgaaggaaag actccgtctt gtcggccctc cagaccccc 240
 tacctgtgag gggtagggaa ccctctggg aggggccctt ccaacaqctq gcttcgttgg 300
 taacagtggc tggatgggtc ttgtcaggtt ctgggtcttc tggagataga ggggtccca 360
 ccagcatctc aagggtcctg gggataaaag gtgtgtcagg atgcccacctt acaccagtc 420
 tggccgtcctt agggggtcctg gggaccctgg ggttgggggtt gggaaacatc acctctgtgg 480
 catctcttgc tgcgtgggg tctgaagaat ggtttttgg gggaaatcagg tgatgttttt 540
 ttaaataatggg cttggaaaaag catctgtggg gccccagata actttgggtt tcataatgg 600
 ggtttatgg gatcgtcaac 619

[0150]

10

<210> 37
 <211> 788
 <212> DNA
 <213> Homo sapiens
 <400> 37

cgtatcatgtt catgtggggg ttctcagtca ttcttttagat caggcttagat ttaggacaca 60
 acattcaqgtt atagctcttq acctgtggctg ctgtatggcc ctggccctggg ctcgtca 120
 gatcaatctt atatgtggat gacagtggcac aatatgtggg gtaaaaccaag gtatcatgaa 180
 aaatatgggg ctcttccttcc aaaccgggtt gttcatgtgg atgacatcaa catacccaa 240
 tcatgttttgg ctgttttacaa ataaaggcatt gcaaaaacca cactggccgt gaaatggcagg 300
 gatcatcaca aaactaaagaa aagttctggg tattcttccaa cttatgtggc gcaacattcg 360
 attggatggg cacgggtgggg gaccggaaag agatagagat ttagccaaaga tatggcccaa 420
 aqccactgtt tttctggact cttatgtggg gattggggaa acaacattcc ccaggcttcta 480
 tgcaaaagggc cttatgtggg acagacttaa atatcatgac atgaaatgtt cttcaatgtt 540
 tgaaattaca aaatcattgg tattatgtgg gatctcttca cttatgtggg tttacaatgtt 600
 aaaaacccac cactatgggtt atatgtggat gatctatgtt cttatgtgg ggggggggtt 660
 tcgggggggtt gatgtggat ttggccgtggg atccctaaag aqacagaga ggcacaaaat 720
 gatggagaag ggcacccatc cttatgtggg tttatgtgg ggggggggtt 780
 gacgtttg 788

[0151]

30

<210> 38
 <211> 481
 <212> DNA
 <213> Homo sapiens
 <400> 38

gggctcaactt atgggttagt attttcttatt gctcaatattt ttctttatgtt caattcacaa 60
 acttgcctttt tatctgggtt agatggcaaa tagacatgtgg tggagaataa ataacatgtc 120
 atggaaaggg cccttgggtt gggataaaag aaccttggatt aaaaatccaa ctcttaataa 180
 cattatataat gacatgtggg gtatgtggc tttaacccgtt ctctctgtt ctgtttccaa 240
 aqatgtcaata attacattttt gtaatgtttttaaaaagggtt atcactatgtt acaagttata 300
 tattatgtt gttttggaaat aatggggat atcttcttca taactttat ccttaatgtt 360
 aqaaatataa ttgttcatca gacaagaaca taaggggggca actggggacac acaatagggtt 420
 cacattggcc tcgtgtggctt acgtgtggctt aggttagggcaaa acqacagaga cggatgtggca 480
 c 481

[0152]

<210> 39
 <211> 192
 <212> DNA
 <213> Homo sapiens
 <400> 39

63 gttgttgtta caagaqcat gaggacttag atttcattt tataaatttt ttaatcattg 60
qcattaaagta gacctgcqaa aatqagaacg aatatatatt aagtaaaatt aagttcaaaa 120
gttagtagta aacaccatgt gacagaagct gggccagtat atatgttaat ttacgtgtt 180
taqqatgcqca ac 192

[0153]

<210> 40
<211> 821
<212> DNA
<213> *Homo sapiens*
<400> 40
cgqacaaaga tgatttat ttctatccaa aqcaactttat aatattctctt qgaaaatcta 60
aaatcattctt gttatccaa cattttata ctatcatcat ttttagaaaaat aaaaggccgt 120
cgttatatac tagaaaaattt ctttcattat atgcaaaata tttatctctt ctatcaaagg 180
agattaaaga acaactgca gaggaaaggaa ggtccctgaaa gtgttcattt tggtatctac 240
ctaccccaac cccaaqacat aaagacagat aaaggcacta agatgttagt atgtggctag 300
tcctttcaat aacccagtc gtcatacag atacccatg ggtatatttc aagccactct 360
ttgagccatc gatggtcattt atttggtag ttcacccaaag gtaaggccat accagctgtt 420
aaaatgtgt agagattaat caacaggctt gccacttgct aatccctcc aaggatgtq 480
tgc当地gggtt ctcattggtc ctgtatqaga atcttgcac tgc当地gggtca 540
tgtccacaaa tactgaggta tagccctgcat gccactaaaa ataacaaagg tttcaaggct 600
ggagacattt tcaaccacac tgtcatgacc atctataqct ccagqactca gtggtaggggg 660
aacaatattt aacggatgtc cctggcaaa qaaaggctgtq aqtaacttcg cctcaaacac 720
ataqatcgc ttatctgcaq cagagatttt ctggcccttc tctgccaqgtt ttttgagggt 780
cttggatqaaq tatatqccaa ctccgtat ttggatcgaa c 821

[0154]

```

<210> 41
<211> 669
<212> DNA
<213> Homo sapiens
<400> 41
gttgcgtacc cctttgacag ttttqtaatq aaggcaaggaa gtqaatttga gcaaggcccc 60
cacgtqccca tqtqacccqa qcaaggcagc accccqcccttq gqcccctttt cccqccctqcg 120
agggagacct qgcgtctgccc ccgcgtctcg cagagtacg ttgcgggccc catcaaaggq 180
cqagqaaagg atqaggaagc gctttqgtt ctgagqcgac ccctctgtca tqtgggtgctc 240
cctattccctc ctccctcgctt tcctcqctt cccccccttc ctgcgcggcc qcaaggcaccc 300
tcgctactca ggagttctgt ccatcctcg tttgtgtggc tgactcaqtt ctgcaggcccc 360
cttcctccca tqcaqaaagg ggacaaqtgtq qcaaccqcc cccaaacctt atttttqgt 420
aattttaaat ggggggttccc agagtgggtq ctggcaccacc aqgggacacc acacagacca 480
ccaggctqca ggaggqgagac acaggccgtq ctgtctgtt tatggagtgc accctttaaa 540
agttccqaggq tqgtctaqcc aqggggccqgtt qgctcacqcc tqtaatccca qcaacttqgg 600
aggccgaggc qagttcaggaa gatcgagact attctggctt qgctagcqgt qtgaaacccc 660
gtctctacg 699

```

[0155]

<210> 42
<211> 398
<212> DNA
<213> Homo sapiens
<400> 42
gatctqactg ctgttqqqat gaqagactga aataqaagag tcaaaqttqa cctqqaagtt 60
ataaccctgt ttaaccqqa qaqttAACAT tgacattqat aqaqtaaqq gtqaaCTTAq 120

65

66

atggggagcc cattcataat tgattttgt ttttcactt qcatgataq tagtcaaaca 180
 atataaaaa qattcagtq aaagtctctc tctcgatct tctctatct atgcccagt 240
 ccctcttgg agaggatcac ttttaccagt gtctgtctaca tccttcgtq gaggctgtaa 300
 qctagggtc ggccaaactgt gacccatggg ccacatccag cccactacct aatttgcgt 360
 gacccttaag ctaqaatgg gtttatatt tggtgccc 398

[0156]

<210> 43
 <211> 424
 <212> DNA
 <213> Homo sapiens
 <220>
 <221> variation
 <222> 2
 <223> n is a or c or g or t.
 <400> 43
 cnaatagacc ttgtgattt tactgctqcc cacqctgctc caactcaqg cccactggaa 60
 cctaaatgaa atccccatc ctaaccacgt qcaactqgccc tttgccatct ctctcacc 120
 aacttcttcc aatccccatt ttattcatta tctccagcca tggactt tcttctac 180
 ctttcttcta tggatattt atctctctca qgcctttqca ttgttcttct tctgcgtgg 240
 aatgttttga agatccctqgc taaaqcaq accttcqca caatcatc tggcatgtta 300
 ctcgagtttta tcttgcac agccctatga gggcttattt gtgttactq tctatttac 360
 cccattagaa tataagttcc aqgggttat cttactqact gctgtacccc cgtgcgtgt 420
 tcga 424

[0157]

<210> 44
 <211> 788
 <212> DNA
 <213> Homo sapiens
 <400> 44
 gttcttagttt aatctqagca atctttcac ttttgcgtq gatgtatattt aqcaqgtaa 60
 tactcctctt qgcttaagac ttccatctt cagtttaat tgggttctt ttcqcatqtc 120
 ttgcaaaat ttcaaggctt ttggcaaaat tttcacactq tcggaaqtc aqctccqaaa 180
 tggccgcaat ggtctgttt ctqacttq atctcttq ccaatqcaac tttcctcqca 240
 aagacaaccc acqgtatagt qaaactqctq ctttagcctc gtgttcttgc aactttctq 300
 aqggtttqaa cttctcaaaa qacaaggqgc aqggtagaca qctqctqcaq tqaactttq 360
 tagagaggaa actcccgqgt ctggqgctca ttcccqcggt tcttcqgtcc gtcctttta 420
 aqgtcqccac tccgggttctt gcttttaagc ctcttcacgg caacgtctc aqccctqac 480
 qccccatatq cccggcqctq atgaaaggaa cqaatqggat qccqcaqct accctqacaa 540
 gtgattattt ttatqacgt tttqacgaga agaaaattqg qtcgggccc qagtcqatac 600
 qatcacaqgt tqggcqggqgc aqgattcga acccacatct aqggacqgc tctqggcctq 660
 actgggtccgc accqcggtccc tgcacqgtt aqgagggtc qacqccctcca qggccqacgt 720
 tccagagctt cggccccqaaac cqcccccqgc gtacgggqgt caccctqgaat gtctaatqcc 780
 gacgtttt 788

[0158]

<210> 45
 <211> 831
 <212> DNA
 <213> Homo sapiens
 <400> 45
 tttttttttt cccccaggaga gggggagacag gaaatctgtt tctgtactt ttcataagtc 60

accagcaaat attttaact cttttttct ctcacatctc aaatccaatc aatttccaag 120
 acctttagat tccacaaaa cagttctct tcacatcacatc accttcttc caatccctct 180
 acaatttatca agatccctt tcttcttcag taacccacc caccataac ttcttctcat 240
 gccaggcttc tcccttactc aatctatccct ataatttgc ccaagttatt aggttacaat 300
 tccctcatgt gtaacatgga gataaaaaatt actgtacagt actgtattga gataaaaca 360
 tgattatatt aaccatctaa catagttctt aacatatagt aggaactcaa taaatgtcac 420
 ttctctttt cttaactctct gtqaaatcaa aacagtqacc tggcattcag gggctctgc 480
 aatctaccca gttcccttagc tctgtccatc aaatattccc ctatgttatg ggtcacgtt 540
 ccaaccaaac tcatactgtq tgcaagtattc tttqacttgg cctggcattt ttatgcagac 600
 agatcactgc tcatgtcttc catccttaca tttqacttgg cctttccctc catctccgc 660
 ctaagtctta ctgaccctca aqggttattt atctcaataq ctgctgctc agaaggcctt 720
 tcttaagtgc ttcatacaag tgagaccttc ccctctctt caatcccctg aaaccaatgg 780
 actacgtga cagagccctt tggcattgtc ataaatatgt gctgtatcg a 831

[0159]

<210> 46

<211> 866

<212> DNA

<213> Homo sapiens

<220>

<221> variation

<222> 859

<223> n is a or c or g or t.

<400> 46

tcqatacaqg qcaaactqag ttqatctqca tgagagagtt aacaggcatt tgggggctgg 60
 aqctgacag accaagagacc gggaaaqgt cggaaaqgt tagcaqcaq gtaqagggc 120
 aqaaagagag cacaagtcaq atcctqgta qaaagctccc atagtggcaa atqaccaaqg 180
 cqtttcttgc acqtaatgc ctcatgtcca qccttaatt qctaccttgc ttatqctgtq 240
 gtgtcgctt ttgttaccat caaacctcca ataaaactgc ataacacqgc cgggtgcgg 300
 ggcacgccc tgtaatccca gcactttggg agqctqagga gggcagatca cttcaqgtca 360
 gggatcaqg accaqqctgg ccqccatqgc qaaaccctgt ctctactaa atqaccaat 420
 tagctgggtg tggtggcaca tgcctgtatg cccagctact caggaggctg aqgcggqaga 480
 attqcttaaa qccqqaqgc aqaaqgtggca gtqggccqag atcqcgcac tacattccq 540
 cttggacaac agqcaagac tccatctcaa aacaaaagca aaggctgggt qcagtggtc 600
 acqccctgtaa tctcataact ttqgggtggc qaaqccggc qatcacatga ggtcaqgqgt 660
 tcqaaaccqg cttggccatc atqgtqaaac cccttctcta ctaaaaatac aaaaattag 720
 ctgggcatgg tggtgcacgc ctgtatgttca agctacttgc qagactqagq caagagaatt 780
 qcttqaaactt qggqgtcaqg qtttqcggtq aqccqagggtt gtqccattgc actccqcc 840
 gggcaacaqg aqcgagatnc catccg 866

[0160]

<210> 47

<211> 495

<212> DNA

<213> Homo sapiens

<400> 47

cgaaatttgc qcgatttatt gtaaacagga taacccttgc actccaccat actggctgtq 60
 tcacatgctc cctttttcac atcaactqgt caaaaaaqaqaa gactttctta agaaggctca 120
 gqtaactaat aacttgcattc cacaatctc atttactaag cactqctaa actqcaqgt 180
 ctggggtagt ctgttattgc tcatggactg cagtttggac ccatcttaat gaatccatta 240
 gtatqctgtq gqatcaaca qcaqctctq atccaaacctq qaaattqgaa aatqctqacq 300
 cataccaaac aattqcaqgt tcatggaaqaa ctacatqaga ttgtctcag atcttctcat 360

69

70

gtgtcattct cataaaaaaag acacttaat ttcaccatt tttaaaaaa gagctaaaag 420
 ttactcaata gcacataatt ttccacataa aaggtaggtg gattttatt tactagtaca 480
 gtggtagatc gcaac 495

[0 1 6 1]

<210> 48

<211> 360

<212> DNA

<213> Homo sapiens

<400> 48

cgggacgtgc aqtgacccca cttggcagca gacactcatt tctgatattt ttgtatgcca 60
 aqtcttgggt aaaaacaacta aqtgatctct taaggaccca gggtcccttt ttgtccctgt 120
 tccttgc(ccc tcaccaccac tttttccatg tgccaccctc tcataagaac tcagaagccc 180
 agggtaggtg caaaagggtc ttttaatcc cctagtggca tatgaaattc tggatgtttg 240
 tgaattttcc tggggaaagg gtctatgtgt gccattagat tctggaaagg gtgtgtgacc 300
 tcaaaaaaaq gttaaagaccc actggaccga gtccttttta aatgqaatgt catggatcaq 360

[0 1 6 2]

<210> 49

<211> 468

<212> DNA

<213> Homo sapiens

<400> 49

ctgatccatg tgatagaaac tagggcagct tatcaatttc atgtggcgta agttaaqagt 60
 gggtgggctg gagctqaagg atccatgtatq gccttgcctt qatqtcgtgt qcctcaactt 120
 qgatgcttca qtcatctggg ggctggcttg aqaaggctgg tcctcattca tgactgtac 180
 acttqatagg aqgagttgtt qgtagccagc tttgtagcca aactaccaga gctcttaaac 240
 aaatcgtaca tcagggagaa agcttaaattc ttgatattaa atqactgttt tgggcattac 300
 tttcagaata attattagaa gtgccttgat taaactaga taactaactgt taaaatatac 360
 gaaatcgtat agtaaaqgtgt ttcttggaa atqgttqaag taaccqgtt tqatattatq 420
 tttgataqgt gcaaaaattt aqatttcgccc agggggcggtg gtcacqg 468

[0 1 6 3]

30

<210> 50

<211> 404

<212> DNA

<213> Homo sapiens

<400> 50

cggcagagct ctgttacaga tatgaaaaca aaatgtgaca caatgaaagaa gctqcataaa 60
 aatactgaaa cctaaggatctgtatgata taggatgtac cctggcttac catagaaaag 120
 tcttttagata qtaataqca ctaagtctaa aqcaaqgag ataagatttc atqttggata 180
 ctgttcaag aactgttattt qqaatattaa agggaaatcc tatgttagta agttaaaaaa 240
 accaacatc atataaaaaaa taaaattttt gtagtataaa acctagcttt tcttattttt 300
 taaaaggaca tcaaactata taattgtca agagcaaagt ttgtgtttt qgtgagctgt 360
 atttcacgtt acttgttagca attatatttq gtgacatgga tcag 404

[0 1 6 4]

<210> 51

<211> 731

<212> DNA

<213> Homo sapiens

<400> 51

ctgatccatg accgggtaaa aaatgtcttg aaatcacatc aggtcatca aagacattt 60
 tatgtgaga gagacctaaga gaaagatgtatq gaaactggac tggaaaagaag tctttatgt 120

71	72
agagactctg cctatgacag tgaccccgag tatagcactg ttcacagggag caaagtcttq 180	
gaaccaaccc aatgtccat caacgataga ctggatatacg aaatgttqaa tgaccatata 240	
gtcctcatct qaaagtcttq qcaaggatqgt ctgacttcacq cqagtqgttt tqagataagt 300	
tgtgttatac ctcaagtttg gtaaagtctg gagatgtaact aacttaaaaaa gtatgtatgt 360	
ttacatagaa tctaaattgg aaaagacaat aatatcagca tttagaaagt tgcccttaag 420	
aaaaaaaaat tattagccat aataqgtttt acatttctcg taatagaata tcctaaqca 480	
ataaaacctc cactatgtga actctaagtt ctgaatatcc gaattttca aatagttctq 540	
agtttatattt aatttgttct ttactgacta gaacacttcc taatttatac catcattttt 600	
aacctcctca aacatactgat atatattttt tccaggttta ttqaggtata attqacaaac 660	
aaaaattgtg tatgtaaagggt gtacaatgtq atgttttqat atgtgaatac atttgatgt 720	
tctggcttcg q	731

731

[0165]

<210> 52
<211> 676
<212> DNA
<213> Homo sapiens
<400> 52
cgccccgggt tctttatccc tgcgttgcgatgtcgcc tcgcgtgcgc gacccctcgc 60
atgcacqctg cccgcataqcc acqtcgcacq ctqtaqccac agacccttq ctcggggcctc 120
acccaaaggcc aaactccaaa cacaatcaga accagccaa gaqacttc ctgggcacqg 180
ccaccqctc tccccctcc agtgtggcc ggctctqca gggtccqagg gctqcatctc 240
taccaqccag cccaggqctc ttccccgggt ctgcattca aqggcaatta cattttaaaa 300
aaaaaaacq aaaaaggta atcacaaaac caaccctcac ttcacagggt ctgtaagtca 360
ctcatqaaac ttgcgtcttc ccqagacqagq gtcccttccc caqctcaggc acaacagaq 420
ctggcaggct ctggcaccct gggccttccc cggqagcctc ccatttggc agtggagcca 480
taaacqggga tccgagaaga qagtatccac tttttttta cagaagaqg qgactcacag 540
cataaaacqgg qgtqgggggg atcctqattt tqaaaataat ctatgttaq ctctcttct 600
atcaaaaacca acacatccctc ttcttgc caatccctc cccacqggga cacctctcg 660
gttcqggacc aatcca 676

[616]

10

<210> 55
<211> 280
<212> DNA
<213> Homo sapiens
<400> 53
cgaatgtagg ttgttttatt gtaaatgtca tcaaaatcca gaacagcaga aacatattaa 60
tcagttgaa attttagaaa tccttagca cttgaaaaag agtattacaa atgcatttat 120
atcacataga aagtcaqcgta atacaaacta qacaqcgq acatagttct ttctggcat 180
tccaggataa taagaatatt tatcaattaa aqgtcaata tctgtcttcc tggaaaataact 240
ccaaacctga gtcaacacac attctttcq gattggttcc 280

280

〔0167〕

<210> 54
<211> 282
<212> DNA
<213> *Homo sapiens*
<400> 54
ggcttaqtcaa gtqaaqcaac aqqaatqtaq aagaacaaat aaatctgtaa qtagctqtqa 60
tcaatttagtt qtaaaacaqca ccgcacttgg accaqqctgt ggtcacccctt aacatctatc 120
caaaaqctga tgaaqaqcaq cagctgaaagg catcacqtgg tacaagagga qcaggtaggc 180
ctgaaaqagc accccacccc aacttccctt ttctttctta ttctctacta tttctcacat 240

75

76

gcaatcaatc atgacgacag agattcaacc aaatctaagt taactgtcga agaaaaggtaa 180
 agatqgtttt attaqtctca ttaaaactagt aactgcaact aqagtatcc tcagtattct 240
 aaaaaaaaaa ggttc 255

[0172]

<210> 59
 <211> 403
 <212> DNA
 <213> Homo sapiens
 <400> 59
 aatgttcatt actttaaaag cttacattt aaaaaactcaa gctttqcttt tctataatq 60
 ttaacttttc catattatca taagaatcaa ctttctgatt aatatttctc cttttttct 120
 taagaagtca atagttcttc atatccacaa tcatgqatca ttttqagga catgccccaa 180
 ttccaaatgt aactgtqagg caggaaaagt acactcccaag ggaagtcaga gtaagcctgt 240
 ttccaccqca gcacagcagt gagcacagct aggcaaaatt ccacgcaagt gcaaatcagc 300
 atcaqagaga qaqttcccaq ggcagacaga qcaactcaqaa tqcttccaac caqqctqcta 360
 tgcaccaaaa qcttggtaa ccttttctct qtggcgattq atc 403

[0173]

<210> 60
 <211> 523
 <212> DNA
 <213> Homo sapiens
 <400> 60
 cgttagatgt qqqtttcaact tttttggccaa qctggctca aactccctqag ctcaaggat 60
 ccccccqccct cggccctccca cagttactggg attacagggtq tgagccactq tgcccaqcc 120
 aaagggttttta attacagaaaa atctcaatct tctqattttt aaaaatgtqttt 180
 tggtaaaaaatgt aagaaaaggaa cagagaaata cttttaggaag ctattgtqac agtccagaat 240
 qagataataaa ttttqaactac aqttcttqat atactttggg tatctgtccc caccaagtc 300
 tcatgttqaa tttagaatctc tagtqctgaa qgtggggcct qgtggggaggq gctqcatct 360
 ggattatqggg qgcagatccc tcatggcttq tgqcaqctt cacgataqgt agttttgtq 420
 aqatctggtc atttqataqgt atgtggcaca ttccccqccc cgccttactc ctqctttcac 480
 catgtqatgt qcttqctccc cctttqccct ccactatqac cga 523

[0174]

<210> 61
 <211> 465
 <212> DNA
 <213> Homo sapiens
 <400> 61
 cgaggaaaga tgggggtttca ccatgttggc caggatqgtc tcgatcttt gaccttgtqa 60
 tccacccacc tggggccccc aaagtqctqg gattacaggc gtgagccacc ggcctqgtq 120
 taagtctgtt tttgtattct aaatqaaatc tttaggacact qattaaaatt tagtaaaaagg 180
 ctttgaggag qgttccaaatq gggatccaga actcgctaat gctctttaca aagctgtacat 240
 caggaaatgt qgcatacaca ctaqgttta actatqtaat aatqtaaaaat tatttqaaat 300
 qgattaaatgt cactaaatgtt qaaqcaaaacc aatataqttt aatqgggttt tttatatttt 360
 cattaaatattt ttaaagggttt tattqatcat qtttcatqtt qcttcttqta cttgtqatgt 420
 qcatqgqagq taaaqtttqgt qgagtaatca aqgtqctatq accqa 465

[0175]

<210> 62
 <211> 417
 <212> DNA
 <213> Homo sapiens

77

78

<400> 62

cgttcaaaqt agctaaccatt gccttggtc tttataaaa tgattctgaa ttgttggaa 60
 attctgttt ttttcttctt gttgtgttt tggtttttt agatagggtc tcaccctgtc 120
 acccaaggctg gagtacaatg gcacgatccc agcttactgc aacctccgcc tcccaagggtc 180
 aaqcgatcct cccacccatg tctctcgatt gcaagggtgtqa gccactgtgc cttgtctgaa 240
 tgctactctt aataaaaaac aaatactatg tttactaacc acattcattc ttgggttcac 300
 tgatccctt ttgtccact agagtctacc ttctgattct ccccttccaa gcaaaqgaa 360
 tcttgcattt ttactgtcag tcttgcattt gatgtatqg tactggccta tgaccga 417

[0176]

<210> 63

<211> 403

<212> DNA

<213> Homo sapiens

<400> 63

tcggtcataq qctctctgca gggggaaqca catcacacgc tgggtggctca ttctggcagt 60
 ccaacacctgc attgtcttta cacaatccgt catgcaactt tgatttcaca ataatcaggaa 120
 gcatatttcattt ttttattccg tagcaatagt ttcagggggt ctgcctacaa tgattattgt 180
 aaaaattcaa ataaaaacaaa atttttatct tagaaactgaa aqgtccctt gtcatctgaa 240
 ccttaaaagac aactattgtt aatatattat tatttcaaaa attgtggcc agcagtggct 300
 cacgcctgtg atcccaqcac ttggggaaqgc tgggtggcc qgatcaactt gggcgggaa 360
 ttgcagacca ggcctggccaa catggtaaaa ccccgctctt acg 403

[0177]

<210> 64

<211> 400

<212> DNA

<213> Homo sapiens

<400> 64

cgtcaataqg atttctttcc atttatggta caaacatgt cattaaatac atttataact 60
 ctttccataa aggtcctata tttcctgggt cacattactg taagggtctca tattttcatt 120
 gtttttgtga aqgtatctc tatttttaa gttgtatctt ttaattttt tatgtatggaa 180
 aatttcaac atacacaaaa gtagaqgaaqaa cagtacaatg accccttagt accatcagca 240
 gattcacagc ttatcaattt agagccagtc tcttctattt aqgggtggcc gataaaatac 300
 atgttcttgc ttctggggcc cctactcctt tctctaaqgtt tatctgtctt cctttgttgc 360
 octagttcaa ctctgtctca atcagtaccc ctatgaccga 400

[0178]

<210> 65

<211> 390

<212> DNA

<213> Homo sapiens

<400> 65

tcggtcataq caaaqaaqaa tgggtggatq taqgacagaa qtatattttg tacataqaca 60
 atcaaaaaca agataaaaatg catgagtatg attacaqgc tgggtggctca cgcctgtat 120
 cccagcactt tgggtggccq aggcaagggtqa atcacctgaa qtcaggagtt cqagaccq 180
 ctgactaaca tgggtggaaacc cgggttctac taaaaataca aaaaatttagc cagatgtgg 240
 aggccacqcc tgggtggccq gctacttggg aggctgaaqgc aggaaattq cttgtatccq 300
 gaaagcgag qttgtggatqaa gccaagatq caccattqca ctccagttq ggcacaacaaga 360
 qcgaaactcc qtcctaaaaa aqaaaaccq 390

[0179]

<210> 66

<211> 307

79

80

<212> DNA

<213> *Homo sapiens*

<400> 66

tcggatcatatgg aattttqaga aaataaaqca tgctgtctt aggaattttt 60
atacttctt gtctttcttc cttaatattt gtttcttagct gctttggca atgatqaaatt 120
gttatgtatg cattaatgtt ttqcaqccca aaagtgttc acattttcc tatataaagat 180
ctgtggatgtq tqtgtttcaa aqaaqaaact acagaaatgt taaagcqaa aaacctgaaat 240
gtgatgtgca cattttcatc ccacatggac aatgtatgtq ttttaataaa tggaaatttc 300
aatttcq 307

307

[0180]

10

<210> 67

<211> 678

<212> DNA

213

<220>

<221> variation

<222> 651, 674

223 n

<400> 67
tcggctcatag aqtgaagaag ttaatggta aataatctga cactgataac caqtatgaaa 60
ataataqtgt taacagtact ttaaaaatgg cagtqgaagc aacccaagtq tccatggatg 120
qatgaatggaa ttaacaaaat gtggatqta catataqtgaa aatattatgg aqcccttaaqq 180
aaggagtctg acataatqct acagtatgaa tgaccttqaa qatactattc taagtqaaat 240
aagccagtca caaaaqtata aatactatat gatttcattt atatqaggtt cctatctata 300
qaagtcaaat tcagagqaaa qtaaaqtgaa atqgtaqtg ctaqqqqttq aqgqagaqq 360
qgttaggaaat ttgtatttga agagaacaaa gtttcatttq aggaagatgaa aaaaattctg 420
qggatggatq qttgttagttq cacaacagtq taaaaatattt taatacccttta qaaactgtcac 480
ttaaaatgggaaa taaaatggta aattttatgc tatgtatattt catggccagt aaaaagtaata 540
qaaaaaaactg cagttacttt tacaccaacc ctaataaagca caataaaaaq aattttaaaq 600
qggctqgtgt qggtqgtcac acctatagtc ccagcacttt qggagqgcttta ngtqggagga 660
tttgttggccaaat ccangaaat 678

678

[0181]

<210> 68

211 873

<212> DNA

<213> *Homo sapiens*

<400> 68

cgttagagaca qggctccact atgtcaccca aqctggtctc caactcctgg gctcaagcaa 60
ttatctcgtc tcaqccccc aaagtqctga gattacagat gtgaaqtacc atgcccagcc 120
ctataacaaa aqtttagaqa tattagtqat cttctaaatc ctatgataaa ccccccttqaa 180
caaatacaata aaatacccaa tatgaaactgt tgccqagtttt aaattgcatt ttcaaaatgc 240
agaaaactgtt acgtttaaq aaggaaatgt tctgcaaccc ccttttatca cttgactttt 300
tatatttggc agacaqtccc agtaggcaaa gtctggcttc cagaacacag actgtggca 360
tttggatctc ttttatcaaa qgaaqggcaat qgatgtccat tacaatttag attqgaattt 420
cctacaggat tctttgcaac tccttcctgc ttcccaqcta gctttgaggq aqagcttagc 480
tagaaaqtqa aattttqqa tgqaattctt atctacaagc aagataaaaag aqtaattaqg 540
taqtccttc atattttaqt qggqacataq aaaatqacaa gtcatttqta aatcaaaaa 600
caaaaataaa tatqcacagq ataatqqaac ttgaatctac ctgacaataat qaaggcacct 660
tqtttqata aattttqaga taacttaatc cagtaccatq tqctgataaa tcttqgataa 720
aqagacttq qggqacataq aqgctccqta tttcttaatq ctacatqaca tcttattqaa 780

81

82

gcttttggtc ttggaaatt gtggccaatt ttatttcta caatggcca cacatgtatct 840
 ccaatccat agattttat gactatgac cga 873

[0182]

<210> 69

<211> 620

<212> DNA

<213> Homo sapiens

<400> 69

tcgggtcata ggcatgttgc agtacttgat agcaggaaat tacctggaaat aatgtcaga 60
 gccatcttgc atgactgtgg agttagctgg ttttattata ctctgggtg acattccag 120
 ctaataaaat taqaattaaa attataatct taqaqcaaq cctctggaa atatgtatt 180
 aaccttttgg aatttaacaa gattgtgtg tacctccatg acctatattt aatgtctt 240
 aacctaagcat ttqaattgtat gccaagatgt gcaagacata gtctctgccc tcaaaggcta 300
 ttaagagaca qactatatgt acacagattt ttttaattata tttaggttta attcataaaa 360
 accctqaagaaa aaaaccttagg caataccatt caggccattg gcatggcaaa ggacttcattg 420
 actaaaacgc caaaagcaat gcaacaaaaa gccaaaataq acaaatggaa tctaattaaa 480
 ctaaagggtt tctgcacggc agaagaaact accatcagaq tgaacaggca acctacggaa 540
 tqggaggaaa tttttgcattt ctgtcatct gacaaaggac taatatccatg aatccacaaa 600
 gaaacttaaag aaatttaccq 620

[0183]

20

<210> 70

<211> 545

<212> DNA

<213> Homo sapiens

<400> 70

aggagtgcata gggctttatt acaaattggaaat ttgactgcta gagaqqccct tctccatct 60
 ttcttcgtt ccttccccc tcccaaagac atccctctag gggagggtcag tagggcatta 120
 ggttggggaa aatctggaga gtggaaaaggq gcttgcctt tgcataagtc ctctgaaaca 180
 accactggat ctggaggctg gctccaggatg agaatcttct aatgtggaaatg gtttagctct 240
 catcttcaag gtccatttctt tctacatctt gggggggctt tgccttttgcctttga 300
 gctgtgggtt actagtccctg gctggcttq aqggggcttc cacttccatg qctgtcttct 360
 ctttctgggc aqggccgatc tgctggggaa gttttctgcg ctttccctt gacagtgtaa 420
 tggggcactg tgactggac gcccgcctt tgggggggtt ccggcgtgatc agcccttgg 480
 ctatcacagc cccgaccacc cgggtgcctca gacqccgcctc ccgactcaac acccqccqac 540
 gtttq 545

[0184]

<210> 71

<211> 575

<212> DNA

<213> Homo sapiens

<400> 71

cgggtctaaat gatcgttttta ttgaaattttt gtaatgtaaat ttgagatcat gaaacatata 60
 atcttaattca aaagaagaat cataggatacc aattcaacac cattaggact actttcacaa 120
 tagtgcacc ccaacttataa cactttttt tgggtctctg actgtgttcaaa aqattgtggaa 180
 aggttgggtt aggttggggcc taggcattac ggttaagaaa taacagatgc taatttttttta 240
 gtggggaaatg gaaactcatg tttagggctt ttcataatgt ctgttgcata aataaattaa 300
 atgttggaaatg aqggccatgc tattgttcaaa cagaatatacat aaaaatacat tctgttgcattc 360
 acctataaaac aaaagatgtt aqggcaatgt gtcaatgtt ccataqccgg taaagcatatg 420
 aaaaaaaaaatc tagtgcataatg agttgtttagt tttcaacaac ataaaaatgaa atgtttcttag 480
 gatgttggaaatg attgttgcattc ttaacattt gatgtttagt gatccataac attttgcataatg 540

83

84

aatatgcaaa atgcccctc cccaactatg accga

575

[0185]

<210> 72

<211> 546

<212> DNA

<213> Homo sapiens

<400> 72

tcggtcataq cactggaaaq aaagaqcaq ccttaattct cctaacattq ccgtttagat 60
 tagtttggaga caaattcatt ttccattaaq tcttttattc attccacaaa tatttcatga 120
 gcatcaaatq ttatgtccaa agcaccatataaqtgttggaa gttgtggcaaa agtaaataaaa 180
 taaaataatac cgaaatgata ccaaaaacaa accccccacc accaacaaaa accaaagcaq 240
 aacaacccca ttataataq catcaaaaq aatacttaq aataaactta accaaggagg 300
 tgaaaaactt gtacaaaaaq tataaaaatq tgcataaaaq aatcaaaqgtt accaataaat 360
 gggaaagatac cctatgtaca tagattggaa aatttaatac ttttttttttcaatactac 420
 ccaaaaacaat ctataaatct atgttaatctt gatcaaaatc tcagtggcaaa ttttggcaaa 480
 atagaaaaaq ctattctaaa gttcatgtgg aatctcaagg aaccctgtt aqtcggcaaca 540
 atctcg

546

[0186]

<210> 73

<211> 516

<212> DNA

<213> Homo sapiens

<400> 73

cgcaagatg ggggttccacc atgttggcca ggctggcttc gaactcctgaa cctctagcta 60
 ggttttaatt actaactggc ataactacca tgagcctctc ccacgtttct aqcaactaaaa 120
 aactccagat tttccctta gaaaactact caaaaaaaa gtttagctaa ttatgtttcc 180
 cttgttggaa aqtaacaaa tgcaatgtttt atattttttt tagttctggg ggttttcgtt 240
 ctatccctg ccaatgctg gacctggggg tggagcaqgg gtaaggqaga aqccaccccg 300
 acagctggaa gtacctggct gcccacggcg aggctgtgtt gttccctggg gacacaaac 360
 aaqttcacaq ggttaagatc agatqagggc tcaccatgac tgcgtggaa caggatgaaa 420
 gcaatcaqc ggcacatggctt gaaacagggat ttctttttt aaaaggqctc tatatqacca 480
 caaaaatgaa caaatgtccc tctcccttat gaccga

516

[0187]

<210> 74

<211> 452

<212> DNA

<213> Homo sapiens

<400> 74

cgacttggaa aattggtttta ttctaaaaaa ggggtgacatt tataagtgg gctagataaq 60
 taacaacata gaaaaattat ttccctccgtt aacaqaqct aaaacatgtc aaacatataca 120
 catacaattc ttcttattata tagctatcca gttggaaattt ttttttaaca cttatgtaca 180
 gggcactgtt ctaaccattt atgtgaatca atttacttaa tcattaataa taacatgagg 240
 aatatgtttc aataatgcaq aataccggaa aattatagtt ttagcataaaa aaaatttttt 300
 tttttggagac ggggttccccc tctgttccccc aggctgtgtt gcaatggcggtt gatctcggt 360
 cactgcaagc tccgcctcccc ggggtttatq ccatttccctt ggcctcaqctt cccaaatgtac 420
 tgggactaca gacatgcacc accttatgacc ga

452

[0188]

<210> 75

<211> 419

<212> DNA

85

86

<213> Homo sapiens

<400> 75

tcggtcataq cacccaaata tgggggtt tataatgggtt tataatctata gataactagtt 60
 tagaaattaa agctgagaat ttaaaaaata cattggagtt tttgggattg ctttttaatg 120
 aattaaataa gccaataagt attaacataa ataacatctt ctgaaaaatg acctttcaaa 180
 aacaaaataa tgagtggcat tggtttacat tttggacat cacttcgtgt ctgttaataq 240
 gagatagtgg gttttcatgt ctcccttqaa ttcaatgtgt tgccggagttt attcaatgat 300
 ccttagtgg qgcatqgtgg agtqgtqtcgt tagtcccaa tactcagaa gtaatqgtg 360
 gatcgcttga gcccaggagt tcaagagcaq cctgggtgac ataatqagac cccatctcg 419

[0189]

10

<210> 76

<211> 390

<212> DNA

<213> Homo sapiens

<400> 76

cgttagatg atgttggcca ggctggcttc aaactccctgg gctcaaqcca tctgaccgccc 60
 tcagcctccc aaagtqctgg gattacaggt gtaagccacc acgcccacca caagtqgact 120
 tatttttaac agattcccaa tgccctccagg ataaaatccc aaattctaaq cctgtgactc 180
 cagatatcac acqagctgat ctttctggctt cttccctta caattgaatt tattttttt 240
 gactaaqag tccttcatqa ataqcttcca ctttcccttc atgtctqtc aqgtgatttc 300
 accttccctt tccttccctt tttctctaca ttatctaa ttccttctt cttctgggc 360
 accgttcaaa ctggatctgc ctatgaccga 390

[0190]

<210> 77

<211> 321

<212> DNA

<213> Homo sapiens

<400> 77

cgcttattgt ttctgttggat atgggggtctc accacattgc ccaqgctgga 60
 ggcttcattt ctaatcttc tatccccctt tcattggctt tgatgatacc ctgtqcttc 120
 gggtccctqa ggcccttcag ggccccacat taatqagqca qccacccatgc aqgtctqaca 180
 tggtctgtcaq tgacccactg ttccagtctc tctccatgtt qcattccacc caacggaaag 240
 cagaagacac cgtccggcag ttggaggtgc tctccacata aqgtgacttc ccctgggtgc 300
 tcccaqccac cctatqaccg a 321

[0191]

<210> 78

<211> 319

<212> DNA

<213> Homo sapiens

<400> 78

cgttagggca gggttttqcc atgttggcca gggtggcttc gaaatccctqa cctcagggtga 60
 tccacccacc tcagcctccc aaaatgttgg gattataaagc gtgagccacc acgcccggcc 120
 ttatggta attatttcc ttacggatgc attacttact gcctttata ctttgggtca 180
 gcttttttaa aaggttaaag taattactgt atctcaattt caaaagggggg gcaqgggtgg 240
 ttgtgaaagg aaggatgtta tggcttqac atttgggtttt ataaacttac atattgtttg 300
 gattcttac tatgaccga 319

[0192]

<210> 79

<211> 772

<212> DNA

87

88

<213> Homo sapiens

<400> 79

cqcaagatq aagacacccat atttaatggc aaaataqtga tggaaagcaq caqaagtaaa 60
 accctcattt tccctataac agacaaaatcg cttgaaaatc cagtttaata aaatccatg 120
 atccctctga acaccagtt cctgatcctc atctctaccg atggcttcata tcataattc 180
 ttccatctc atttcctgtt ttttcgttt ctccatattc tccaaattgt gttttccca 240
 gctctctgg ctctctctg cccacttctg aaagtttctt tagttcttctt ggttctgtc 300
 tgctgaatc tcctqagat ctttqctcta attctccca aggcctccata taaatttctc 360
 ttgactctct tctgcccac tcttagtagt ctttgcctgt ttctccatgt ctttgcctca 420
 attctctga caggcccttg gcaactgcctc ccaaattcgct tctggctttt cctgcctcca 480
 cactctgtt tccgcttqctt gggctctqca aqcttccccca aaaagctctca tcgtccatcat 540
 ctgtctctt ctcatctca ctatggtccc actgcttcac caccctctg ggtctatctc 600
 tcaactctctt catgtccttc tctctgttta ttattctatc ttcatacagt gaaqttctc 660
 cctcgctttt ttctttqaga ttctccacgt agttggttt cttcagttctt ctqatgatac 720
 cttgttggc atcgtccaaq gcaactgtca tggcctqctg cgcctcggtt ta 772

[0193]

<210> 80

<211> 474

<212> DNA

<213> Homo sapiens

<400> 80

acaccacgccc cagccatctc ctctagaatt cttttcatgt cttttgaatc tcctctctct 60
 ctcttcata ttcttgcctt ttttttqttt tttatcttctt tatcatctcc tctgatgtttt 120
 gaaacaatcc aacattgtct gttattcac caatttggca ttttgcagat tttaaattgg 180
 ctgttactqc attcactgtg tttaaaatcc tggaaactatq ttttgcactt ctaagctatc 240
 tttctgttac acaaaatggtt ccctttttaga actgctqaca aattaccatq atacaacatc 300
 ttacaaaaac tttaaqccaa aattacaat taaaatgtt ctaagttttt ctttgcctgtt 360
 ttttgcagca gttttaaacaa gggacttta ttctqatttt tcaagttqgt cccttgcctt 420
 tataacagct ggttgcattt aatgtttqg tqaatcttctt gtttgcctgt ttttgcctgt 474

[0194]

30

<210> 81

<211> 277

<212> DNA

<213> Homo sapiens

<400> 81

gttagagatgg ggttttgccttattttggcatttggctgg gctgggttgg ccaggctggat ctcqaaactcc 60
 tgacccatcaag tgatccgccc accttggccg cccaaatgc tggattata ggcgtgagct 120
 actgqccqca qccctggttca ttatcttgc acataaaatcc aaaagaaggcc caccctqca 180
 cgatctccaa gggccctgtt gggaaatgttcc tggaaaggag aacagctcac 240
 ctctqccatq atctcgatct ttcgttaccg acgttttq 277

[0195]

<210> 82

<211> 305

<212> DNA

<213> Homo sapiens

<400> 82

ggtqatqcaq atttcaacaq taactctggaa aactgtqaa aatgttatt taaaatata 60
 tatgtatataq ctactgcaca gtttcaaaga tggattcat aaataatgtt ggttgcactq 120
 attaattttta taacaattac tggacttcca agttqatqca aacacqcaq gactcataact 180
 caatattaaq cactaataat atccctcaqg cgtactacaaq ttttqatqtaa gtttqatqta 240

89

90

acatatatat tttaaatgt atgcattat acaaactgt tatattatgt atgggtgtc 300
agatc 305

[0196]

<210> 83

<211> 457

<212> DNA

<213> Homo sapiens

<400> 83

gatctgacac tggaaaggaa cccactcgca tggatgtgt gaaaggact ttatgtta 60
tcctgaaaag ggcctacaga ggccgtaca aaactctaa acaaaggaa gttgactctg 120
ctttgaaaaa aagggtggcat gccagcacta gctaaaaaaq gaaatgtcta gcccggcggt 180
gtggctcagc cctgcaatcc cagcactttg ggaggccgag gcccgttggat cacgggtca 240
ggagtttgg accaagctgg ccaacatgtt gaaacccgtt atctactaaa aatacgaaaa 300
aaaaattatgc cgggtgtgtt ggcgggaaacc tgtaatccca gctattcaagc aggcggaaacc 360
agggaaatcg cttgaaacctg ggaaggcaga gtttgcgtg agccaagatc gacccactgc 420
actccagcct gggcaacaag agcggactc catctcg 457

[0197]

<210> 84

<211> 222

<212> DNA

<213> Homo sapiens

<400> 84

cgcctttt acccatcgat ttccacagct accatgtgtt cttttagaat catactaatg 60
tcgcggatc tggcaaaacc atggatataat tctcacacgc agatgtgtatg tcaaatctgt 120
ctatctctg ccccaaact ctccacataa tgaacacttt tcctctgtcaa acatgtat 180
tggtttggatg aggaactcaa gtcattatgtt aggtgtcaga tc 222

[0198]

<210> 85

<211> 217

<212> DNA

<213> Homo sapiens

<400> 85

gatctgacac taattctatg gctggagcac atcccttaac ctctccgaga ctccat 60
tttattataa aatttagaatt tggaccatg ggcctctaaa tgcgttccat ctttgcgtct 120
ttgtctttc tgcgtacata aacatcgat ctaatcttg ttagtggcat tattttgaaa 180
acatggaca gaaacatagg aacaatattc tttatct 217

[0199]

<210> 86

<211> 740

<212> DNA

<213> Homo sapiens

<400> 86

cgtaaacaat agaaatttat tggccgacca tctcggcctc ccaaagtgtt aggattacaq 60
gcgtggccca ccgtaccccg ctggatgtt acgtttatc tttatgtt caattataaa 120
attattcacc cacttagaac agaaagcaat atggaaaaaq gaggaaagaa aggtcttctg 180
tacttggatg gggccagggtt gggccaaaggaa actggcactg tggagcttggatgtt 240
caggaccgc aggttgcgtt gcaactgtaca tggcccaaggaa aatgttgcgtt 300
cttgatgtg agtggatgtt gggccaaaggaa tggccaaacttc ttaagtattt ctctaaaaaca 360
tccacacttc aggttgcgtt tggccagggtt gtttgcgtt gggccaaaggaa tggccaaacttc 420
agcatcacac aatataatcca tggccagggtt gtttgcgtt gggccaaaggaa tggccaaacttc 480

91

gtcaaggccgt cgctttatgg ctcgactgtt tttatctgc tcctctaaca tcttataggg 540
 gttgaaggaa agcggttct cgtggccac aaagtttcc cagcagtca caaactctga 600
 gagggaaagag gaaaaccagc caqqccccag qggcagcagc tgcagtgcca ggaattctcc 660
 ctcccttcc cctctttct cttccctcc ctggacattg ctctgtccgg ggaattctcc 720
 ctcccgggtg gtgtcagatc 740

92

[0200]

<210> 87
 <211> 327
 <212> DNA
 <213> Homo sapiens
 <400> 87
 agggtaagat aggtatgg accacatctg agatagttga cctgggtctt atggaaaccc 60
 ccatccgtcc accagtctca tcgtccatgg atggatttca aggagttggc aaactaaacc 120
 tggaccaccc tggaaacaca tgcacttttc taagcaaggag gttcaqagac aatatacaaa 180
 ttttaaagaa gtctgtgata aaaaattatt qacaacaac aaatgaataa aattaaata 240
 tactgcctaa taaataccta ggatcaagta ttttaggat taaatgtgg tatgttaacaa 300
 gaaaagctcc ttggccagtg tcagatc 327

[0201]

<210> 88
 <211> 503
 <212> DNA
 <213> Homo sapiens
 <400> 88
 gatctgacac tggaaactaac actcataaaa agtggacacag tactttcgt ctggacccaa 60
 gtggattaaat ttccctggtaa aacaaacaaa caaacaaca tttttcaat gattttaac 120
 aggattcaaa gtcgtccacaac atattaaaaaa atccaggata caattcaata ttactcaata 180
 tcaaaacaac ctggaaaatc taaacagtac tcaaaaggaaa agacaattca ataggcagta 240
 acactgaaat gacttagatc gtggaaattat tagacaataa aagtaggtat gaccattctc 300
 cactcttggaa attagaaatg tcttggaaattt agggaaagata gaaatctca gcaagaaact 360
 agaaaactgca gaaataatcaq atggaaattt tagacccgaa aagatgatat gggcccgat 420
 gggaaataaa tataacaqg aaaaqagtca gtggacctqg aqgttagatca atggaaatg 480
 tacacaatgaa atatacattt acg 503

[0202]

<210> 89
 <211> 714
 <212> DNA
 <213> Homo sapiens
 <400> 89
 gatctgacac caaaatcaat gccttattaaa gaaaaaaatc atagtttggat ctttacccaa 60
 attggaaataat ttgccttcc aaaatgtcacc attaaagaaaa ataaaaatac aggacacatc 120
 gggatatact gcaatggggat ttttagaaatg ccacaataaa gtggacatgt caatggaaatc 180
 agtcacatga atttctgggt tcccgatgtc tataaaatgtt acgttttacat gacatgtgt 240
 caacaatgtt atgtctaaaa actatgtgtc tggataaaaatgttact gctaaatggct 300
 gctttacacag acgtttggat tggacacaca ctggccggaa aatggccccc acataacttgc 360
 ttggaccagggt tcaccacaaa ccctcaattt gtggaaaaac acatgtatgtc tggatgtgt 420
 taaagccaaatg cacaataaaatgttcc attcctgtgc aqgtggccacagg caaacaatgg 480
 atttcaatgtt cttatgtqatca tggccggat acacaaatggat acataatgtgc 540
 aatgttggat aacaaacaaa caaacaatggat cttatgttgc aqgtggccacagg 600
 atggggcaaga gggccggatca caatggggat gggccatgttcc tggatgtgt 660
 aqcaaqggaa ttggccatgttcc aqgtggccacagg aqaccacggcc tggatgtgtc tacg 714

[0203]

<210> 90
 <211> 499
 <212> DNA
 <213> Homo sapiens
 <400> 90
 cgqagaagga qtttcgtct tattcacag gctggactgc aatggcgtgt tatcaqctca 60
 ccgcaaccc tcgccactcgg gtccaaagaga ttctccgtcc tcagcctccc aagtaqctgg 120
 gattataggc atgtgcccacc atgcccqct aattttttt ttttttttq attttttagta 180
 gagggagttt cttcatgttc atcaggctgg tctcaaactc ccgaccctcaq gtgatccact 240
 cqccctcaqcc taccaaagtq tqgggattac aggaqtqagc caccgtqccct qgtctgacaa 300
 attcatttc tqttttttt qgtctqctct gtttccaaca agagcctact tcattatctc 360
 cccaaatatat ctacctagtc catagactgg ctcagagcaq gcctttgaca atatttaatt 420
 ttatcccaca atggcatttc aattcccttq tcattctqag aaaaatacta acaqcatqgg 480
 aatttqgtgc atggatcaq 499

[0204]

<210> 91
 <211> 499
 <212> DNA
 <213> Homo sapiens
 <400> 91
 cgtgaqcaat aaagctttt aatcacctgg gtgcaggcag actqagtc aaaaaggagt 60
 caqcaaaaggq agatqgggtq gggcagttt ataggatttq ggttqgtagt qggaaattac 120
 qgccaaggq qggttattctc tagcqqqcaq qqqcqqqggt cacaaggqtc tctqgtqggg 180
 agctcctgag tctaattqtc caggagaagg aatgtcacaq gqctaattq a tcaqtaagg 240
 tqgggqagga acaaattcaca gtqgtqgaat gtcatcaqtt aqgcaggaa ttqgttattt 300
 tcactcttt tqgtqgtctt cagttqcttc agqccatctg qatqatataq tgcaggcttq 360
 qgctcagaga cctqacagct ccctaatttq cttctaataa ttctaaggaa qaaggaaqag 420
 ctgtqagaat tqtcatttqag aggaatcatt ctctctqag taaatqagac aaatcttctq 480
 ttttqgtqtc atggatcaa 499

[0205]

<210> 92
 <211> 485
 <212> DNA
 <213> Homo sapiens
 <400> 92
 cgaggtqctt tcttcattt qagttttqag aqatcttta ctattctqaa ttcaqggccct 60
 ttatcagatt tqtgatttqa aaataacatt tcccaqactq tcacttggcc tttcattctc 120
 ctaaqaqtat ctcttqaqa acagaqttc ctaattttqa tqaqgtctaa attataattt 180
 ttaaaaatgt qgatttqctt ttqcttaaaac caaaggcaca gtttctctta tqtttcttc 240
 taqggcttct atagttttag gttttatatt cagatataq atccacttq aqtaatttt 300
 tacatatqgt gtqagatatq gttcaaaqtt tatattctq a tataqatatttq attqgattgt 360
 tacagaacca ttgttqcaa agqctqtcct tttctcaqtc cattatctt qcatqttq 420
 qgaaaaccat qqaactqtc ttqcatqta qtagaaaatt aattqgctq cstatqcatqg 480
 atcaq 485

[0206]

<210> 93
 <211> 600
 <212> DNA
 <213> Homo sapiens

95

96

<400> 93

cgttagagacg gggtttcacc gtgttagcca agatggtctc gatctcctga cctcgtgatc 60
 ccccccgtcg qccctccaa aqgtqctqqqa ttacaqqcqg qaqccaccgc qcccccqccqg 120
 cttacattct tttaaagtca qaagaaaaaa atqaaactttt aqgtcqagaa aatattttg 180
 ttatataata ttqatacgtc atgtttatata caatqcaqtt qtaaaatatq ttaaatatgg 240
 atacatataat ttaacttaca tqtgtgtqaa atatatgtgt gtgtgtatq tacaagccac 300
 aaqgcaaaag tgctcqggtt ccaagaaagt caaaacqtc cccqccaccc qaaaaqccq 360
 aaacccctcg qggcqgtat tatqgcaqg caagttqaa qatqgcaqgt qgtqacaaaa 420
 ttgtqaaaaa tatcacattt ctctcaccct gtcaaaaggca accccaaagt qaaqtcaaaa 480
 tagaqatqaa atqcaqggc acttcctagc tqcggtatct tqgacaqag aqttgaccaa 540
 cctqgcccctc agtttccca tqgatacaat qqqqgatqgq ttqcccttcc qaacqtttgc 600

[0207]

<210> 94

<211> 306

<212> DNA

<213> Homo sapiens

<400> 94

ggtcagctgg agaaaqttat qtttttctaa qaatttqtc acatcatata aatttcaaac 60
 gtattgcccct atagttatttc attagttttt cacttaacat ttqaatgtct acaqaaaaag 120
 tqatatqttt ccqggttcct qccaaqattt tcaaggttgg ctttacttc tctqagtata 180
 gtaagcatgg ttqctttata gtatacqccc qgttaattccca atatttgaag aatctqcgaa 240
 gtccatttctt ttttttctta gtcacttaca ttcatggaat ctcatatctt catgtqcatg 300
 gatcag 306

[0208]

<210> 95

<211> 295

<212> DNA

<213> Homo sapiens

<400> 95

cgttagagaca gggtttcacc atgtcqgcca qgctqgtccc qaactcctga cctcaggtaa 60
 tccccccaccc cqgctccca aqgtqctqgg attacagqcg tqgqccatcg cqcccgatcc 120
 tatttcaattt ctttttagtq tqtcaataqg taaaatqgt ttttttaat atqtaqggt 180
 tatttcagac tatataatata tttcttttgc ctcttttacc ctactqggc caqaggaqgt 240
 tcattatcta gaattacagt taqaaaagaa qtcqccatq aactqcatgg atcag 295

[0209]

<210> 96

<211> 295

<212> DNA

<213> Homo sapiens

<400> 96

cgtttcaag qcaaggcacg taatgtggac attatatctt cgtqcaaattt aggattactg 60
 qaaqagtat ttttaattaa aattttaaaga qacataqagg caaaaatgtgt ctgcccattgc 120
 acactatqga tctgtcaata caqaaattt qttqaacaag qctaattgtct qaaagcacca 180
 tqcaagttttt caqcccccgtg attacatttq ttttctcaag aqgtqgtttt tataatctac 240
 accctqgqgt tcccaqtttq taaaactqtaa qctttaccct tqtqacatqg atcag 295

[0210]

<210> 97

<211> 254

<212> DNA

<213> Homo sapiens

97	98
<400> 97	
<pre> acaatggctt ttgtacattt ttcctctcg ataaacctca gaaatggatc catcatacct 60 tctggcagat gaccgtgtct cctttttaga gtaaatggaa gttgccagac aqaaaattcc 120 ttgattttctt tccataaaaa tqaqaaattt acccacatgg atgcccctttt tgggggggtct 180 ctcctgtat tgcattggaa gaaatggccc tcttcattt ccaaggctaa tccctccacc 240 tgtgcattggaa tcaat 254 </pre>	

[0211]

[0212]

{0213}

```
<210> 100
<211> 386
<212> DNA
<213> Homo sapiens
<400> 100
cgttagatg gggcttctcc atqttggcca ggctggcttt gaactcctga ccttaaggta 60
tccacctgcc tcagccccc aaactgctgg gattacaggc attgagccac agcgcctggc 120
ctacagttaa ctttttcaat ttttggatgg gactcgatt atttctgtcaa gagttgtgtca 180
gttttgcatac tgctttatata agaactacat ctctgtaaac acaagcataat taaaaatca 240
```

99
 taatattgtt ttccttaaga aagaaaaact ttaataaga ctgatcaqca aqaatcacaa 300
 catcaaataq gactgattaa tattctctac cagtagttcc taaacctggg tagctttta 360
 aacatataca ctactggacc aatcca 386

[0214]

<210> 101
 <211> 610
 <212> DNA
 <213> Homo sapiens
 <400> 101
 cgaattttaa attttatta aaaatcacag tttttaaaaa tgcgttgcgtt cacaatgttag 60
 ctcttctqaa atqatqaa aacacaacac actgcttaacc aqgtggggaa gctctctqca 120
 gggctggggac acctgacaat qcaqctgttag cctctqcaag taaaatcca ggcgacttg 180
 caqtcattgg actqatgtcc aqgtqcaatc accatacagc agtacaggc agggctggct 240
 gataggggat atggggagaag gacacgctca gatgaaaaca tgcgttgcac gattttcacc 300
 actgaacaca ctgtttctq tgatagaaac tgctggccct gctggggac aagatattca 360
 cggcctcact aqccqctgtq atqccaccq gcaqccctgc ccctqatqct cttttgttac 420
 ctgctaaaga aggaccataa ggtaaaaggc accttaccc atggagtqag cccagacccc 480
 aqggaaaaggc ttqggtagaa caatccaagg ggcgacttg atgtgagaat ccagcccaag 540
 ctgctgctc taqaaqccctg gaggcccttc ccagcttccc aaatgtqcc ctttcacat 600
 gaccaatcca

[0215]

<210> 102
 <211> 479
 <212> DNA
 <213> Homo sapiens
 <400> 102
 caaacqtcgg gqctacatq gatctgtgt tqaatqcggt ctgaaatqat cctatqgatt 60
 ttcccggctg gttgccacta ctgtacaaca ttcaqgtgcc acatccatct gtgccattaa 120
 gctttttgtq gacatqagag atgccttcc cctgctgtat gacatqatt tqaqaaqttq 180
 gaaqaaaatq acaaatacq ggaqaaaaca tccaqgttc ttacctgtq atqaaatcq 240
 ccctcacttq gtgcttatta ccqgttattt aagaacaata acaacaaca aattqatqaa 300
 catccaagaa qcacatatta ggaccaaaga tagcatcaac tgcattqaa gqaactgtq 360
 ttqgcatt ttatqacatt tttataaagt actqtaattc tttcattqag gggctatq 420
 atggagacag actaactcat tttgttattt qcattaaaat tttttqgt ctctqttcg 479

[0216]

<210> 103
 <211> 279
 <212> DNA
 <213> Homo sapiens
 <400> 103
 tggattqgtc cagtggttggaaag tggagagcaaa caagagqaa tggactccc aqccqcaatq 60
 aqgatttqca gqgtggaaqgc tgccctcaac gtatcatqaa gcccggcgcac catggqaaac 120
 actttcaaaa qtcqataata aatatctcaa atqacatat tttattqac cttaaaaatq 180
 catacaaaaag tgcatttqat tcgtttaaa tacaatttac ttaactctt ttaacgcca 240
 ggattaaqcc aattctactc cataaggctt ttgataccg 279

[0217]

<210> 104
 <211> 254
 <212> DNA
 <213> Homo sapiens

103

104

<400> 108
 ctqcttgatq accqgtqaaa catttacttc tcccaccatq ggctgagacct tggatgtgtc 60
 ctttacctct ctctttgtca cttctgtccct qaactqccca tqagtgacca qqqqcttccq 120
 acagaagtgt ctcacaatcc aaccacacag cagaaaggca ttttccttq gatggagct 180
 ttgtccctgc actgactttt tqtgtcatct cctttcatt caatgtgaga tactcacctt 240
 gattcataga atgttttctc attccagggtt tqtattttt taatcaccaa aagaagcact 300
 tattcataaaa qactqgtatq atgtataact gtgtatttag tttgaaaatc ttcaccttat 360
 agaagcaata aacaatattt aggtgtattt tcctgtqaaa tttttataaa tataagactt 420
 tattacg 427

[0222]

10

<210> 109
 <211> 326
 <212> DNA
 <213> Homo sapiens
 <400> 109
 cgtacaggat aatgggttgg ctgggtatat atttaataat tctatatgca ggctcaaggc 60
 attttccctt aaccctctga agacattatt ccataggctt ctgttggttc tgccctttg 120
 tgctttctc ttcccttca gagttctgca attgcaatgt taggtcttct ttttatttaa 180
 cctgcttggc actcactgtg tttcctgaat ttgagggttc atgttcttca acaattcaag 240
 aaaaaatccc agccattacc cctcaaatat tatctttccc caattctatc tcttcctca 300
 ttttgaatt tggatgcaat qcgtac 326

[0223]

<210> 110
 <211> 415
 <212> DNA
 <213> Homo sapiens
 <400> 110
 gatcgatttgcataccctcg qggtaacaagt cagtttctac actgggtgtct gatcccttgc 60
 cgcgcctttt ctaccatctc acgggtggggta tggccgcaagg gctgtgcacc cagaagaaga 120
 tggctgttagt cgtttcgccc ttccataatqg cctgagactt tctttctgtt qgtgtctgg 180
 agctgtccgg ctgggtggccc cctatttgc catttaqcgaa caaaccacag qattttaaa 240
 aacaaaaaca tcccaagatt ttttcatttc aaaatgcttc aaagtccaca ttagatcaga 300
 tactccgtg tcggcacatt cagctgagggt tcattacaat cgagactgca atgtgatcta 360
 tggatgttcatct tggatgttata ataaaaagct tcaggaaaca agccaaagc cctcg 415

[0224]

<210> 111
 <211> 567
 <212> DNA
 <213> Homo sapiens
 <220>
 <221> variation
 <222> 4
 <223> n is a or c or g or t.
 <400> 111
 cqgnaggagt cttactcttt ctcccaggct ggagtgcagt ggcatgatac cagctcactg 60
 caatctccac ctcccagggtt caagtgattt ccctgcctca gcctcctgag tagctggat 120
 cacaqgtqcc tggccaccacq cccggctaat ttttqattt ttagtqagaq caqagtttca 180
 ccatgttggc aaggctggtc tcaaactct gacatcaagc aqttcgccc cttcaagcctc 240
 ccaaagtgct qggattacaq gagtgagcca ccggcacctgg cctgtaccat tatttqatca 300
 acatctgtct tcccccttaq actggaaagct ccagggtggc aqggaccatg tttgtccctca 360

105 ccaaggatcg caaggccctgg acacggtagg tactttgtaa atatttggat actagatgg 420
tggacacgaaa atqaatggat ggatgggccc taagctaagt tctgtacatt tttaaaaaaa 480
aatcagattt ccaccccaact aatqactaaa cttggactg tagtttctgt gtccctctgt 540
tttgcccttg cacatggcaa tgccatc 567

[0225]

<210> 112
<211> 574
<212> DNA
<213> Homo sapiens
<400> 112
cggtatttaa agaaatcaa aacatttca tagcccccta aqtgttatgg gcttggacac 60
tgtacctcgc ccaggtaaga tcactcgtgg gtaaaacat gaggctctca cccgtaaaggc 120
aggatttta tagaaggaag gttagtcttt caacctatgt cctccttctg ttccacaaag 180
tggaaagcca caagccctac aaaagccctg caagtcccaq aqgctgcagc cgtatttatt 240
cttcaaggcca agactctcag gacagagaqc acccatgcac cccgcaggct qcaggccatc 300
tccctcgcatt tggactgtc ctqaggatgg cggcttcatt tttgtccctc ctacccctg 360
gtactccccca ttcccttatgt gaaqctgcgc cttaaaccgc accccactgtt tttagtccgtt 420
ttcaggcagc tgataaaagac acacccaaaga ctggcaatt tacaaaqaa agaggtttaa 480
tggactcaca gttccacatq gctggggqagg cctcacaatc atgggtggaaag gctaaaggca 540
cgtctcagcaggcagat aagacaatgc gatc 574

[0226]

```
<210> 113
<211> 283
<212> DNA
<213> Homo sapiens
<400> 113
gatcaagtcc cactatcaaq gataagccaq taaqtgttgtt ttagcataaaa atagtttctt 60
gcttttgccct ccattcatct gttatatttg aaaaacccat agacaacttg atgtgtttgg 120
ggtacttaat tgacatgttt gtaaqcagac taaqtgtgtct tgttgtgtaaq ttcttaacat 180
tgacattctg tggatatttgcacatctcaatt tgacatcgaat tttttgggta gttgtgtttat 240
aatttgttaaa tctttatattt taaaagattt cacccgttgg cat 283
```

[0227]

<210> 114
<211> 596
<212> DNA
<213> Homo sapiens
<400> 114
caaacgtcgg agataaaagggt gtagatgcac cagtgtaaacq tcttgccttq tttcccccattt 60
tagtataaag atgggtgtatg gcatgtctqca ttccagatcccc catgggtgtat tttggggattt 120
ggggttttggggg atttgtgtgtat gaaaggccaggaa aqgattgtcccc aqcaactctc atttqaatttq 180
aagggttgttggaaag cgggttctcgg gcacagtatg ccaaggccca gggcgtccaa gggacagagc 240
aqaqatcaca gatgtgtaaac qgggtggggaa gqagacttagt qcaggqccatt tctgtatqaga 300
ctttgatttq ggagggtacgaa aqttgtatg aggaatcttq ataaactgtt tagagtgtat 360
ggagatqagt aggaatcacc accattcaca tactggcttq gcaagccact caqcacagct 420
ctgtcaccttgc tcttatggaa gagaagacaa gcaagaqaga atgttctgtat ctqagaqagt 480
tcgtgtaaag qcatgtaaaggc aacagcactt gagatgtgtt qcacatqagc aaaatttcag 540
tttatgttggaa atttaaaggatt aacagattaa aqgataaaat tcttcccttq qtaccc 596

[0228]

<210> 115

107

108

<212> DNA

<213> Homo sapiens

<400> 115

gggcttatca gatgcattgt ggtgatgaac tcattctgt ctctggagga ctgccttca 60
 tacaccacat tgccctgtac aactgctctg catttgagaa tcatacttaa atgtttacgt 120
 ttcaattgcc tggaaaggcagg gccagtgct tatgtcacct ctgacccttt acggaaagggg 180
 acaagtgagg tagtagctca cacctgtataa cccagtgctt tggggaggccg aggtgggagg 240
 atcatttgag gccaaggagtt caaaagaccag caccttgcac ctqaagaccc tctctcctqa 300
 tcatccagca atccaggctg ttccggctt gggctccgtca gtattctgca ctctttacct 360
 ttttttttaa tcaqacacaga gaaaggcaat taccctcaca gtacccgatg aaataagtat 420
 ttttaggtt gtagtqattac tttagtqattt ccctgtatac aaaactgggt tctttttta 480
 aggaagattc acatttctca atctatttt ctttcaaaag agagaaatttta qactttatg 540
 ttttaagtat acacattctg ctttctccc ttcaattttaaaaatttcaaaatggcc 600
 ggtatgcagtq gctaaacacctt gtaatccaa cattttggaa ggccaaggccg ggccgggtcac 660
 gaggctcggaa qatggaaqacc atttcggca actggtaaaa tttcatctt actaaaaata 720
 caaaaatttag ctagacatgg tggcgctcac ctqaaatccc agtactcaq gaggctgagg 780
 caaggaggtc gcttgaaccc gggagtcaga ggttgcagtq agtgcgatc atgtgtggca 840
 acagagcaag actccggctc c 861

[0 2 2 9]

<210> 116

<211> 289

<212> DNA

<213> Homo sapiens

<400> 116

cggaaatttqa agtgcgtcgca ttataatttta ctgggtcagc atccctaaat ccagaaccca 60
 aaatctgaaa tgccctataa aqcatctcct tttaatqctg ccqgactca agaagttca 120
 gattttgaag ctcactggat ttcatttttt ggattttggaa tgctcaacct acatatcaca 180
 acttgctata ttcccaaaaa ccacttacca gttggaaaa aattttttta tatctgtcat 240
 ttctttttc cttctggaa aaagagtqaa atttctatgg tctqagatc 289

[0 2 3 0]

30

<210> 117

<211> 363

<212> DNA

<213> Homo sapiens

<400> 117

gatcaagtcc acaaggcat ggtgcccgg gaaaaacttc cccctgaaaq ggcgcacaagc 60
 aggcggaaacc cagccaataa tgtaatacata tttaggtcgca accagaaggat acattgcaaa 120
 cccactgctt gaaaaggtaat aaattcatat tgtagctataa atgtgttgg tqaattqata 180
 atgaatttat attaaqtqaa tcaattaaqt aattctqaaat tggcttttga aagtctgcga 240
 gaggctggcc accatggctc acacctgtaa tcctaqact ttggggaggcc aqggccaaq 300
 gtttgcgttga gctcaggaaat ttgagaccag cctggggcaac atggcaaaac tccatctcca 360
 ccq 363

[0 2 3 1]

<210> 118

<211> 752

<212> DNA

<213> Homo sapiens

<400> 118

cgactctctt ccgcaacctc cagtcctccc cacaacatac ctggcaagt acacttacat 60
 gaactaactt agataatacq aaccacttaa acccaqgaa gaaactcgtc catcttcatt 120

109

acctgaatta tctggcatgt tgatattctg ttacttatgg aqagcaqact aqgtgtgaaa 180
 agcacacccc agcaacctca gacagatagt cttaaatctat tgccctggat tgtaactac 240
 ctcttagata gttctagata gttccttctt gcaagaaaggq gatgtcaacc tgaggatca 300
 ggttcgtat agttcttatg taatactctc tccttggat gtaaagatgg tgcactcaa 360
 ccaataccag cacacatact ttaagtattt ttctactaat ggcgtatgt ttaattaaac 420
 atccctccac atgcctgcct cgtcccttctt gtagcaatca ggtatctca cggggctggc 480
 tggttagt ctttggacata ggtgtggat aqggtgggtt catgtggatg tgcgtgggg 540
 ccaatctccctt ctttgcacc acccaaggca acccaagggtc tggacaagat ggttccatg 600
 gtttccatg tctgactttt ctggatgtac ctttgcactcc tgaaaccata gatattctt 660
 tagaagcaaa tttttatgtt cgtcattca tgacaagaaa tgaccttca gaaagaatc 720
 tgatataatt ggttttcat acctgcqaaa ac 752

[0232]

<210> 119
 <211> 282
 <212> DNA
 <213> Homo sapiens
 <400> 119

gctccctttc ttttatcttq tctctttca tctgaataact gtgtatgtca cgagagtatc 60
 ttatgtttc ttcaqctgt tgcaatctgaccatgtt ttctgtgtcac tgcattttc 120
 tactccggag aqaggtctt ttggacaga tgctgcccccc ttttagactcc agtgcgtggc 180
 actgccttctt qagtttttc acttccaaat tcttattatc ctgcaccccc accacacagt 240
 caatgactttt aqgttctctt ggttttccatc atggatcaca ac 282

[0233]

<210> 120
 <211> 303
 <212> DNA
 <213> Homo sapiens
 <400> 120

gatcaagtcc cagtagactc attagttaaq tggtacqgtg gtcaagtat gttgtttgt 60
 gtgtctgtttt tctttaggtt ggttcaggta ccctgaaata aatcctaact tataatgtctg 120
 ttgtctcatc atctctgtgt tggtggaaatq gatattaaq tggtttttctt ttaaqaagga 180
 aqagccagg qcatagtgcc tcgtatgtt aatcctaact ctttgggggg ctggggcagg 240
 aggattgtttt cagtagccgg gttgtggactg gcttggggaa catatgttca 300
 acg 303

[0234]

<210> 121
 <211> 883
 <212> DNA
 <213> Homo sapiens
 <400> 121

tcgatacagg cacaaaqqa aqtcagctca acccaatctt catggaaataq gqaagtgtag 60
 taagggggaa atgaaaggaaq acttcctggaa gaaaggaaat cctgaaattgaa gtttgcggaa 120
 gaaataaaagc atttgcggg aagataatgtt gaaaaaaggt gttccaaaggq aqgagataagg 180
 attttcaaaq acattggat gaaaggaaacaa ggttaatgtt aqgaaatatac tgggttcaa 240
 ttcacctgag actcaagaca gcaaataaaac ttctgaaaga tgagattaga caaataggca 300
 aqgttttagat aatggaaat ctttcatgag tttgaaactttt atttattttt gatgtttta 360
 gtgtttttttt tttcaaataa ttttaaaggaaq tgaggatgtgatggataat cagataaaaa 420
 ctgattataaa atgtttctgg caacactgtt gagaaggaaat ttcgtatqag tgcactqga 480
 gaaaggaaqaa tgggttaggg agtagtccag gaaqtaatcc aqggggaaq aataaacaac 540
 gatggccatq gacaaggaga aqatgtcatq tggatgtgactg ttaaaggatqg taqaaactaat 600

111

ttaatgacca aqtaatgtq aaaactgcaa aagaqqaagt ccagtaaaac ttctqttcc 660
 agtatgagaa agatttgatq qcaccactca ctaagataca qgtatgtcac agattagaga 720
 tqaqcaqqa aqagagatct gctqcaaata tqttaaatgc atccacttgg agaggtccga 780
 tgggtggctq cttatggaca ttgtggatq caggcagag atttaaatta gagatcttca 840
 gctctaattct ttgtgaatcc ttaagaatat gtcttaactt ccc 883

[0235]

<210> 122
 <211> 260
 <212> DNA
 <213> Homo sapiens
 <400> 122
 gtagagatgc attttgccca tattgtccag qctggctcc aactccctggg cttaaqgcgt 60
 ccacccaccc ctagccccc aaatgttggaa attacaggtq tgagccactq caccacccc 120
 acgcccattt ttaacaatgt aagaaactga ggcttgccaga ggttagatca tttgccccaa 180
 atcacatgtq gqatccagaa ctagagcccc cagtcttgg attctcaaca cttttccctc 240
 cacactacag gqacttqatc 260

[0236]

<210> 123
 <211> 964
 <212> DNA
 <213> Homo sapiens
 <400> 123
 qatctgactq cactaaataq acagtttta atctcaattq gacatctqaa gatttctcaa 60
 gctgttggq gcttcagcag aaacaattt ttctgttttq ttgattaaac cagtggctc 120
 caccattqct tatcggccctg caacgagctq ttcttttaa ttaaaatgtq gccaccacga 180
 gaaqgcaatt aacaqctaqt aattactgtq ctcttcagaa cggcaacatt ttqcccttca 240
 ttccaattat ataatggaca agtgggggttq tgatcttcaa attagtqagc tttgtccctc 300
 cttgtttggq gqcaqaaattq gcaaggtaat gatggccctgt qggggcactc tttccctctq 360
 ctgatgtqag gcaagaacgc cttaaaggq gcaagtqagag ctttcagaaga cccqcaqttc 420
 caccctttag ataaagacag agaaaqagtc tttatttca gagggttttq atgaggagga 480
 aaqataqgct tacatgtggc ctccaaacag tqccatgtq tttccctct tcttacccctc 540
 caqcaaaagg tqaccaagtt tcaaaatgt qaaaactqaa agtttcaqat gtgcaagatq 600
 qctaaqttct qqaqatccag tggacagact gatqactata gttaqcaata ttgtattqca 660
 tacttqaaat tttcttaaga atgtatct taaatgttct caccatgtq tgacgtat 720
 qcatacacac aaacacacac tcacacaqaa ttactatgtt aggtqattat atatgtcaac 780
 tagcttgatt gtggtaatca tttcacaata tataatgtata tcaagtatq tacctcaaat 840
 atatacagtt tctatggca aatgtacacc aataaagctg aaataaaccg aaggaaaact 900
 qqaatttagat tatgtqatq tttqgtatq ctttggtaag gqataaaat accacatgtt 960
 ggcg 964

[0237]

40

<210> 124
 <211> 516
 <212> DNA
 <213> Homo sapiens
 <220>
 <221> variation
 <222> 509
 <223> n is a or c or g or t.
 <400> 124
 qqqacttggg tqtaactqqa tcagattttc tggaaattcaa qqqqaaqcc gagatacttc 60

113

cctcacagaa attgttaata tcaatgctta gcttcttgc cagttcctca tcactttca 120
 gttgttcttc catcgctctt cgccctttt ctgcctgtct ttttccctt tttccctct 180
 ctqccaacaa cctctgtatq tattttcac tggctttgtt ttcttctcc tcgctggccc 240
 gtcgctctgc cgccacccctg cttatccctt cttcatattc tcttctcgt tccccagggt 300
 tactqagcaq acqaaactggc tgatgtcat cagccacttc ctctgattct tggccagacg 360
 ctctaaqctt qactccctq ggataqgtttt tttqaattat cgtccacagt tccacgttga 420
 cgaagagaat ttcttcgggt atggtacccg aqtcacqac gataccggc ggacacagaa 480
 ggacacqcat aaactcgccctt tttcgacqng tcacct 516

[0238]

<210> 125
 <211> 297
 <212> DNA
 <213> Homo sapiens
 <400> 125
 cgttctcagg ttccccctt ggctgctctg ctcactcga aqtagcagggt ggagtttcat 60
 tctctgttc ttctacatq gagaqattt cctctqaact tatgtctqgt tctggctttt 120
 ctccctcccc ttcaagcaagc ttgctttgg gaggagtttc ccgggttagaa ttcacagttc 180
 gacqaatcgq catggtqcta tcttctaccc tctaaqaaaaa gagatggcca taatcaaaca 240
 tqcattaaaaa aattctcttt tctctcttc tcaaaggqcc taactgggtq tcaqatc 297

[0239]

20

<210> 126
 <211> 527
 <212> DNA
 <213> Homo sapiens
 <400> 126
 gatcaagtcc cattataqag aatgggttat atggtatttq catataacct acatacatcc 60
 tcctgtqtaat ttaaaatcat ctgaaatqatq ttaatcttatac ctaataaaaat acctacacat 120
 cacttcattt qcatggattt aacatagtgc ttgggttca gtaaattcaa gttttgtttt 180
 ttggqaactt tttggaaattt ttcttttctt qaatattttt qattggcagt tgggtqaagc 240
 catggatgca gaaaccatgg atatggggg ctgactgtac ttacttacat atctatctat 300
 ctttqatqaa caaactcttt qggggaaaaq aqggaaatqatq attqagcaq agaaggcaaa 360
 aacattaca taatattttq attctgttac cgtgtqctca tagattgtttt cattctaatac 420
 aagacctaag aqgtatqat tctgtqactg atttagaatt ctgggtqga aatatcttgg 480
 aaccatgttt ttaataaaat tctgtctctt ggaaagaqtt aacaacg 527

[0240]

<210> 127
 <211> 525
 <212> DNA
 <213> Homo sapiens
 <400> 127
 gatcaagtcc ctttatataa aatggcatq tatttgcataq taacctatq acatcccttc 60
 atatacttta aatcatcttctt aggtgtttt acagtaccata atacaatqta aatgttatq 120
 aatagttqtc atactacatt qgggtttttt gttattttt ttatagttt ttcccccaaa 180
 tcttttcaat cggcagttgg tagaatccac aatggqaa cccatggata agcaqgggttq 240
 actgtattaa agtgtttctt aqgattttca aatcaaacat tctacqgtga tttatattq 300
 ccttgacatc accaaggqat qaaagtctct cacacttctc tccaaatta tctcgtaaag 360
 qcaaaqaaaaa qggaaaaaaac cccataatctt aqgggggtqgt ctaaccataa aaagtaacat 420
 ttctataat tttatcttgc gagttcaagg aatattqat tccctcttattt ccataatqta 480
 tttaaacata aaaatatttca cctttcccccc accttcttcc caccq 525

[0241]

50

115

<210> 128

<211> 507

<212> DNA

<213> Homo sapiens

<400> 128

cgtttggaa gttatttatt tctgaaagtq tcatttatat agtttggatq tccccccaa 60
 ctctcatgtt qaqatqtaat acccaqtgtt qgggggtggg cctaattggg qgttttggg 120
 tcattggagt qggatccct cgtggcttgq ttttgccta aqgttagtqg gtgaggctct 180
 cagaagatct ggtggttaa aqgtgtttca cttccctcat tcttgcctt cttctgttt 240
 tggccatgtq atacgcttgc tttcccttca cattcaagaa qaaaaaqttt aattgtttca 300
 aqagaataaa aatactcagg aatccaactt acaaqqqacq tqaaggacct cttcaaggag 360
 aactacaaac cactgtcaq taaaataaaa taggatacaa acaaattggaa gaacattcca 420
 tgcctcatgq taggaaqaaat caaatcgtq aaaaatggcca tactgccccaa ggttaatttat 480
 agattcaatq ccatccccat caagcag 507

[0242]

<210> 129

<211> 509

<212> DNA

<213> Homo sapiens

<400> 129

ctgcttgatq agaagaaaaga tgggtgtggc atccccgtqg atgggttcct ttcgtttct 60
 acaaqaqtaa taggcccgg tagaaactgc tgggttcctcaq tagttggacc atttataqg 120
 cagattcttc ccactctgg aatttggctac aaaaggtagc aaattgtttt ccaagggttt 180
 tagcgtatc gcttcacc aqagtqataa aaaaatgaaa gctgtggccg aacatgggtgg 240
 cttggccctg tggatcccaqg actttggag qccaaqgtgg qaggattqct tggatccagg 300
 aattcaaaac cagcctggc aacataqtaa qaccatctt ctacaaaaaa tggaaaaaaat 360
 tagccgggtt tgggtggccgca cacctgtggt cccagctact caqgqacta aqgtqagagg 420
 attgtttggg cccagaagtt caqgqctqca gtqagctatq atagtqccac tggatcccaq 480
 cctgtgtgac tggcaagaa cctgtctcg 509

[0243]

30

<210> 130

<211> 501

<212> DNA

<213> Homo sapiens

<400> 130

cgtatcatgq aatcattttt cactttttt tttttggag acagagtctt actctgtcat 60
 ccagggtggq gtaaaatggc gtgatcttg ctcactqca cctctggctc ccagggttcaa 120
 gcaatttcg cgcctcagcc tccccggtag ttcagtagtt gggattacag gtgcgcacqg 180
 ccacccccgg ctaatttttta ttttatttcat ttttagtaga gacagggttt 240
 tggccatgttg cccgggtgg ttcgaaactc ctgqagctcaa gtgatctacc cgcctcagcc 300
 tcccaaaatq ctggaaactac aggcatgaaac caccggcaccc ggcccatttt tcacttctta 360
 agttttttc cctgqaggttt cctatctttt cattagggaa acaataqgqg ccaacagggtc 420
 agcctacaqg aactgttgac aaaaaggact ttagttcaag tcctagttt ggcacttgg 480
 agtatataqg acatcaagca g 501

[0244]

<210> 131

<211> 497

<212> DNA

<213> Homo sapiens

<220>

117

<221> variation

<222> 3

<223> n is a or c or g or t.

<400> 131

ggaggatta cacaagggtt aaatgagata ccacatgtca aaacacttag tatacgac 60
 ataattctaa atagtattta ccataattag ccattattgg ccgaggacac acaaggata 120
 cacgattaaq ctgaaatagt aatacaaacc taqgttatttttggccct aaagacttcc 180
 ccttttqgt tctgtttcat gcacatctga ttgtgaact gaagcctcat ggccaaacca 240
 gcttcagggtg ggtatagaca ccgagcccc acataataat aqctgagcc taataccctc 300
 cccagggtg ttctttgtat ctcctgggtt cctgacatt tcggatcgat gattaatgg 360
 taqtttactt gatccatggt ccaatttctt caaaqtgtca tacgctaaat aactgtatqa 420
 accaaatagt tcttttact tcagtaactt tqtagtact atqgttattt taacaatgt 480
 actgtaccat caaqcag 497

[0 2 4 5]

<210> 132

<211> 413

<212> DNA

<213> Homo sapiens

<400> 132

cgatttctaa actttatggaa aaaagccaat gacaatcaat tqccaaaaaa cgagacgcca 60
 aatcccatct cactccagcq ccatqccctc ttctcaaacc ctcgcccgccttcttccaca 120
 tccaccagca caqttctcq gqgtcctqct ggtctcccaq tqctgttctt aqgtccaca 180
 tqaqqccagg tcagggcgggc ctcggccggc aqgacttctt qtagccggaa aqcacqaggc 240
 qcaggagccca tqgagqcccg gtcctttggaa gacaggata cttaaqggaa ttggccaaacc 300
 qgagaaacgcg tqggagaggt qccccgcact qcggtggtaa qcgagqgtct cagttctctt 360
 cqggcttqtc qgctqggcggt ccacaaggct qgacatctt ttccatcaag cag 413

[0 2 4 6]

<210> 133

<211> 411

<212> DNA

<213> Homo sapiens

<400> 133

ctgcttgatg atgtaataaa ctaacatgtaa aatcttaat aaacagtgtt cattcttaat 60
 ctgtccatca actcggatcata gtcagtatttq ttgtccgtat ttccacagatt aqgacaccag 120
 ccttagggcc qagtqccatq gtcacqccctt qtaatcccaq cactttggaa ggcccaqgq 180
 qgtggatcac cttaqgtcag qaqtttagga taqccctggc catcatqgttq aaacccacc 240
 tctactaaaa atataaaatt agccagggtt qgtggcacgc acctgtatc ccagctgtct 300
 qgagggctgtgaa qagaggagaa tcagttgttac caqggagqgt qaggttgcag tqagctgaga 360
 tcacacacca ctgacttcca qccctggccqaa caaqgtaaa actccatctc q 411

[0 2 4 7]

40

<210> 134

<211> 893

<212> DNA

<213> Homo sapiens

<400> 134

ggcggccggca qtcctactt gtcggccggc tggaaatggc qggcacaatc tcagctacc 60
 acaacctcac qattctctq gtcaggccctc ccgaggatgt qggactacag qcgtgtccca 120
 ccatqcccaq ctaatttttq catttttagt qagagccggg tttcaccata ttggccaaqga 180
 tggtctcgat ctcttgcact tggatccgc ccggccctggc ctcccaaaqt qctqggataa 240
 caqggcgtggg ccactgtacc tggccctttt ttttttgaa acggagttttt attctgtcat 300

119	120
ccaggctgca gtgtggcaat ctcggctca ctgcacaccc cccctccctgg gttcaagcga 360	
ttctcctgcc tcaagcctctt gagaagctga gactacaggc acgtgccatt agqcccaagct 420	
aatttttgta ttttttagtagt agacqqqqtt ttggccatgtt qcccaagggtq gttttqaatt 480	
catgatctca agtgattcgcc acacccctggc ccaaagtgtt gggatttacag gtgtqagcca 540	
ctgtgtccaa ccaagacttc ctaaaactttt atgttacaat accctactga ctcttcaggg 600	
aaaggccacgc acaaattgtga tggatccctgt qccataggc ttatacttcc ccaacctc 660	
gtaccctaaq aaatctgtct gacaaattgtt ttggccattac catccttcac catcacatcc 720	
taagataact acttccaaa qgttagttcag gatttaaacta ctaccgtct taggacaagt 780	
acaagttcag caagagggttta taatggagacc ttatccatgg caqagacaaa gtgtqaaaga 840	
atgttaattca ggcaatgcag ggtgtggatcc cctttatctt ccccatcaq caq 893	

[0248]

<210> 135
<211> 858
<212> DNA
<213> Homo sapiens
<400> 135
ctgcttqatg tcagtttaac aattttatqa accagtttct ttaataqgtgt tqtggaaact 60
cttagccaga qcaatggat aatcttaaaq ttataagaaa cctgtacttq ttcaagagtct 120
cttccatqaa tctccttqaa gacgaaacac tttaggatta tagttgcatg caaaagcttt 180
caggaaagca tcagqataaa aattatctt qgatqacagg actttaaaatq qctatgataa 240
gaattttatta tgcaatcgac aaggaaattt tattttttgt gacatcgta ttttaacata 300
aaaatcaaaa ttatqactqg taacatcata acaggacata tcagattttq agqaatttca 360
taccatttctt agaacactca ataacatacc caggcaaata taactcaagg aaqgttaaac 420
atcattqctt atttqacagt qcttaccqgtg taaatttaac atactqaata aggttagttt 480
acatctctct ttttatqagg agaqagaaca gattctttq agatattcta qqaatcccct 540
qgaaactacc qaagtttagt caaaggtaga aagacctaaat ggaggctggg catqgtggct 600
catacctqta atccccqcat tttqggqaggc caaggcaggc agatqtttgc tqagggcaqqa 660
aqcagcctqga ccaacatqgt gaaaccctgt ctctactaaa ttaaaaaatt agctggqacgt 720
qgtqgtqcat qcctqgtqtc ccqgttactc aqaaggctga qgcaggaaaa tcqctcqaaac 780
ctggggagatq qagggttqtag tgaaccgaaa tcgcaccatt qcattcaagc ctgggcaaca 840
gagcccaqact ccqtcctcg 858

[0249]

<210> 136
<211> 780
<212> DNA
<213> *Homo sapiens*
<400> 136
cgttaga~~q~~acg ~~q~~gg~~ttt~~cacc atgttggcca ggctgg~~t~~tc taactcc~~a~~ga gctca~~a~~gt~~g~~ 60
tccac~~c~~tacc tcggcc~~t~~ccc aaagt~~g~~ct~~g~~ tagt~~at~~aggc at~~q~~agcc~~a~~cc gg~~g~~cc~~g~~gg~~cc~~ 120
tc~~a~~g~~c~~t~~t~~gc~~a~~ ttttatatt aaaata~~a~~ct cacagg~~g~~ct~~g~~ ttttag~~t~~ta~~q~~ caa~~c~~ag~~c~~t~~q~~ 180
gg~~t~~at~~t~~ta~~a~~at agt~~g~~t~~g~~ttt caaattat~~g~~ta~~t~~ qtagt~~g~~ttac ccattt~~g~~ct~~t~~ cttaaaat~~a~~ 240
g~~ct~~ca~~g~~cg~~c~~ gact~~c~~t~~g~~aaa ct~~c~~t~~g~~tc~~c~~t tggag~~t~~ag~~g~~ca tgg~~t~~ttt~~t~~ tq~~c~~act~~c~~at~~c~~ 300
ccggaaataa atgttctaaa cacag~~a~~gt~~c~~ta aagtc~~q~~caaa gaaaca~~a~~aga gg~~q~~gg~~c~~t~~q~~t~~q~~ 360
gaca~~a~~gg~~g~~ag qccaa~~q~~ca~~q~~c~~c~~ ct~~q~~t~~t~~att~~c~~c~~t~~ acataaccat taaagaacaa tt~~t~~q~~c~~gt~~g~~gc~~c~~ 420
tatataaaaat ctaat~~t~~taa aaaat~~t~~taa gaaat~~t~~taa gacaaccatt tcaaaaaaaac 480
at~~t~~taaaaaaa gg~~c~~at~~t~~ttcc tttct~~t~~q~~g~~ taagg~~g~~g~~q~~t c~~t~~ttc~~q~~agg~~c~~ ctt~~t~~tttta 540
aattt~~q~~t~~c~~a aqcaaccat qttcat~~g~~ta aqtaa~~q~~aat tcat~~q~~t~~q~~caa caca~~a~~ag~~c~~g~~a~~ 600
aaattcc~~q~~aa tc~~c~~t~~q~~t~~q~~at~~q~~c~~c~~ gccccac~~cccc~~ tagtataatt actatt~~q~~at~~q~~t~~q~~ gtttctt~~q~~t~~q~~ 660
cttacaacaa gaaaaac~~q~~cc~~c~~ tttcttaata gaagagg~~g~~ct~~c~~ tq~~c~~ct~~g~~gtaa ggg~~q~~ggcc~~cc~~at~~c~~ 720
gg~~q~~gt~~c~~t~~t~~ct~~q~~t~~q~~ q~~c~~q~~g~~t~~t~~q~~q~~g~~q~~qa t~~q~~tac~~a~~g~~g~~tt~~c~~ at~~q~~tt~~c~~at~~q~~c~~c~~ cat~~q~~t~~c~~aa~~q~~ca~~q~~q~~a~~ 780

[0250]

121

<210> 137
<211> 560
<212> DNA
<213> Homo sapiens
<400> 137

ctqcttqatq qcaaaaqqqq aqqtqqaqq gtqqqaaqaaqg aaatqtcqaa qcatqtgctc 60
tqataaaaata aqggcaggqa aacaagctq aattacttqaa attacttqaa ttttcttqtc 120
ttaaaaactq aaaaaaaaatq tagttacacq ttaaaaatctq caaatqgqtt ttacacctct 180
qattnaaca tgaacttata ctaatqtttq gaatcttataq tcagaataat aqcgagctat 240
gtacttqaa taggttttq aqgggagagg tagaacagag agagaattaa qaaggqgatct 300
gactttataaa agactqaaat gtqattqagq tqataqaaaca taccatqtt accaagaaat 360
tgacaagctq ctggctttaa gcttatgcaa gtqgtqtagtq gqaaagtqgg agqgtqggaa 420
gaggqtttqct attttggatt aattcatqca aaatqaaqqa gqaagccctqg tctaaqaqaa 480
tactqtcattt caataqaaat qattctaa ctqctacaga ttaqaqaaatq ataatctqat 540
tgctgttqtt ttgtttqttc 560

[0251]

〔0252〕

<210> 139
<211> 486
<212> DNA
<213> *Homo sapiens*
<400> 139
cggagacaga gttcactct tqtqcggag gctqqaatqc aatqqtqcgq tttcqactca 60
cggttaacctc tgcccccgg gttcaaggcg ttctccctgcc tcaaactcct qagtagctgg 120
gattacaggc atgcaccacc atgcccggct aattttgtat ttttagtaga gacgaggttt 180
ctctatqttq gtcaqactqg ttcqaactc ctqacccctcg qtqatctqcc tgccctcqcc 240
tcccaaagcg ctgggattac ggqcttgagc cactqagccc aacqctgttt qtgattgtta 300
aaggaaqgtac atccctqctct ttcaagtatta tattttctta attqctctta aactttctct 360
tatttttqcc ataaatqaaa acaaaatqca ttattactat tcataatcata caccctccctg 420

123

124

tatccttaggt gacaaggtaa aaagtqactc acacacacat tattttctag tattggcattc 480
aaqcaq 486

[0253]

<210> 140

<211> 397

<212> DNA

<213> Homo sapiens

<400> 140

cgaattaaag agatgaattt attggtaat agataaaatt aacacctatt ttaatatac 60
caaaccctt ccttatataat attaggtaaa taaaaagaaa aaaattatca aqcaatact 120
acaqccaqct acatcqccaa tttacaatq aqtagtaaq taccataaq ttgtttaat 180
atcagggtct tcagagttt tctcaacaaq ttacaaaaga gattagggttc ccaatcagtt 240
cacaaaaqct aatttaaaga atgtagctt actacagttc tgagggtgtc acacacttaa 300
cttgcgtct ctgtttatt tattcatatc tgagggttac tgtttctact aqgatacatt 360
ccgccccacac ccacaccccttq gtcggccat caagcaq 397

[0254]

<210> 141

<211> 397

<212> DNA

<213> Homo sapiens

<400> 141

ctgcttgatg ttaatcatac tggactaag aqgatagaga atqgttaggag ctgggatacc 60
cctaaacatt cacattaaaa caaaaaaaac ccaaagctaa aaaacaactq ggcaggagct 120
aaataaaaaat ctaatttqa qaggctgtat ctggctcagg cctcctactt tgtaacccat 180
ggaatatgtq aagcatttq aaaaactata qcactgatct cacatggca qacacactct 240
caqagqatq tggttggqac catggcqcaq tctggcttagg caqgtggcaqg aqgcqcaqaaq 300
actctgattt ctctccctcgq tcctaagacc qaatgtgtt caqgacatgt ggtcaqggaa 360
qagaqctat ttaactgaac cagtaataq aqgaccq 397

[0255]

<210> 142

<211> 652

<212> DNA

<213> Homo sapiens

<220>

<221> variation

<222> 645

<223> n is a or c or g or t.

<400> 142

cgtctattat ttctttatgc ttactctgcc atttattaaa qgaggaaaaa tqtqaacgat 60
gaaaatattt acttcaaatt ttgacttattt aaactttgga aaactttcc cggagaatga 120
ctttcttggt tgaggtaat ttccctqaagg actqagatq ggttaattat gggcacaca 180
gtgaagtctt aatggcttcc ttaagttcac acctatgaat tcctggatq gccagagaac 240
atqttactac tccaaqgtca atqgtqaaq catqttqatq qaagcatqct gatqaaaaca 300
aqaaaqactt gtggaaatgc tttagttgtc cctcaaatq qaaaaaaatt caattctgtc 360
ccaatcatac tggacaata catatqagac ttcaattqta qaaggtaac aacttcagg 420
aaataaccat qaagtcaaca qgataqtcac cataacctca cataatccat taactqtaat 480
acactattca ccaggcttt qgtttcgtq aqaaaqgtq qgcttqtaqc aaaqqaccac 540
qatacaaaca caaaataqqa actaaaatq ttcttcaaag acatqacatta qaqgtcacaat 600
caqagtqcta aaccaccaaq tcctttat ttgcaagtgg acatnnaagc aq 652

[0256]

125
<210> 143
<211> 659
<212> DNA
<213> *Homo sapiens*
<220>
<221> variation
<222> 6, 648
<223> n is a or c or g or t.
<400> 143
ctgctnaatg ggattggggc caaggagtca gtgcacatcgta gttggtatgg aagatggaa 60
ggagggcaga gcaagtttc ctgttcccta gcctgcqatq ttcaagctgt taggaaqcaa 120
ttcaaggcagg aaagcaagtc tgctccattt aggttagtcc cacatttgcgt gttgtttqa 180
ctttgagcag qacacagaq tccttcaagc cccaccccttc cccctgtgtat atgggcagat 240
tttactcgct catccctaact tcccttagc actccaggg cacaggcata aggtacaagt 300
tttgcaqctg gaaatgcata aatgtqgatc ccagttccag cactttctgg ctgagtagtc 360
tttgatctat cagtcactt ctttgaqgtat caqcttcctc ctctqctatq tcctqggttc 420
catgtttcc ttgcagaqct gcaataaggc tcagagatgc tataqaatgg gtcataqgtc 480
ctagcaggct ggtaaaccaat gctagccata tttcaqgtct aaqattggg caaatgttgg 540
tctgtccctt gttgaccaac ttcttccctt gagaagcaca tagcatctgt acacacagaa 600
gctgtgtqcat gtgaggqatt catacatatc taagaacccctt cttgcaangg aqgttacq 659

[0257]

<210> 144
<211> 564
<212> DNA
<213> Homo sapiens
<400> 144
cqgtttctcc aqqtqatttt aqqtqaqaaa tcagggttqa qagtcaactgc tttaaaggaa 60
aaaaattaag gagcagtacc ttagaagagg atattaaaca ggggtggagt ttqctctaaq 120
acagaqtaga acatattttt gttqctctqa gatacaagga aacctgacta gatacattta 180
agagatctgg ttqagataca gtggctcttt gctcatagta aaaaagtatt aatqaaatat 240
tttaaaattt ggacttccag caactcaqaa tttagtqgtc tatgaaaaag taaacacaaq 300
qaagaccttaq aaaatcaqta tatagaagtt aqgcgtqaa actaaaaqgt ttctttttaa 360
atgaatqaat ttgtgttqat atacttgc tttcttataa ccattattaa agcaattatq 420
aaqgtttat tttaaaaca taaqgtaaaa caactqagct aaaatcatga caactctatc 480
aqcattttctt tatccacqaa tacacatcac caaaaaatgg cactactqat cgtccgtcca 540
qqggaaataa tttagcatcaa qcag 564

[0258]

<210> 145
<211> 337
<212> DNA
<213> *Homo sapiens*
<400> 145
gatcccaqtac ccaqtaaaqt qctctcacat gtgttgataq ctggacatct tccataatqc 60
ttctttcaact aaaacaaaact ctttgcaaat ttaaccctata cattgggtt aaattctaca 120
taaaatcatg qaatttaaga gctggaaagag ggttggcaat aacctaatct agtctctgtt 180
ttaacaaaqa gaaaactaaa tttaaaagat ggtaagtgtt caggccgcgtt qactcatgtcc 240
tgtaatccca gaactttqaq agqccaagggt gqgtggatca ctggaggatca gqaqttcqag 300
accqaqcttgg ccaacatgat gaaaccccat ctctacq . . . 337

[0259]

<210> 146

127

128

<211> 970

<212> DNA

<213> Homo sapiens

<400> 146

cggtaaaaqq aattqaatta qaaaqtcac aatctagaaa cttcaaaatt actqcagcca 60
 cctactactt tccaaaata agcaccatc aaaactagac tctatcattt gttaccatq 120
 cgttctat tactaatgtt ctcaaataa atttqagctt aaaaacagaa aatctacatq 180
 ttcgtgacq acccctcca ccccaacta ttccaccctt cagccccatq gcatgttaaa 240
 gggaaqaaca atgaattctg gttgttacaa qagaatattt ggagcatctq actcttttg 300
 ctgtcaacca qqqtatcaac ttqagggatc cataatqact cqactccctq tttqattcct 360
 tctctctttt tattctctag aaagcccttc ctggaccctc ttgttgact aatqaccctt 420
 tctccacatt cctaaccacc ctgtgcaaac ctctcttga qcacctgcca cattctgtq 480
 aaattaagtc ttatcctccc actgtatttt aaagtctttg aggacqatgc ccatttgcaa 540
 gtcaccqttt tattctctttt aatqatctt aatqatctt aatqatctt aatqatctt 600
 tggaaaggqtc atqatqatqt tgggggggtq tqtgtqagaa qotaaaggac acgttttca 660
 tttgcctta tataqgacc actcatctta acttcaaccc ctattccaa tcttcctctg 720
 aattqataaa atttttaattt qttacattct ctgttagatgg qtttggatga cacagaaac 780
 cccctaaagt catctqcaat qttctcagcc ttqagtttga tggattactt ctgtqagtq 840
 tgggtqagtq tggcagtaga cagctcatcc aatcccttga qttactgttta cttgaqactt 900
 tttcqcttc aqaaaaaaqq aacatqagct aacataqaa aatqcqaaag tcttataatq 960
 gtactqatc 970

[0260]

<210> 147

<211> 317

<212> DNA

<213> Homo sapiens

<400> 147

qatccagtagc qgtcaagcag qcaagatcaq attcacaggt aaaacttqct tqtgtggact 60
 tcaacgtttc ttcctatttc ctgagtggtt ttqatgtgtq agagatagaa tggtaatgg 120
 tggaaaggaaac ttataaagct agtagtccct tttttqgtt qtagaacccc aqtagatataq 180
 aaaaaaaaataaa aaaaatgtqqa tggttqagtq qgctcatacc tggatttctq gactttggg 240
 aqgccaaggc aqgcaqatca cttqaggtca agqgttqag accaqqctq ccaacatqgt 300
 gaaatctgt ctctacq 317

[0261]

<210> 148

<211> 763

<212> DNA

<213> Homo sapiens

<400> 148

cgatqgcaqt atttttaatc ttgaaaagaa aataacaaca cagctcaaqg aqgaaqaat 60
 attctaaagac acaccaaaatc tagtacttct ctgtgtatga ccaagaaact gtttctgag 120
 aqtagqcttag aqaaaaatcc atttactaga accctttctt aqgqcaqtc atqatqatcca 180
 catttgcac tctgtacata qgqgatactq tggtaactqgg aacaaaaact ttttgcac 240
 gtttcaaaac tctttqata ttqctatattt actccctaga gtttagact tggcaaaaaat 300
 gtttagtqattt acagagtcaq agttgaaaat atatctcaca agqaaagaga catccagcca 360
 agaaaaacaca aatattttaaa ccatactgaa aaaaagcata acatqttgag attattacca 420
 caaagataaaa qcccttcttt ataqcccaca qtagtqgqaa aatcaacaaa ttctcttqgg 480
 gaaattatata tccaaagctg accagcagta qgtagtacca tcacacagca attctcttca 540
 aqgttctcca qaatatccaa agaagtcac aqgtqtcacc aatccaaacg acaagqccct 600
 qggcgtatata qgtctattat actatagaca aqgtcatctt taacaccctt qggctttqta 660

129 atcca ttaaattctt ctgagggttt tctccaaaca gttacatctt tcgttttctt 720
rtcgc cacttatctt cttcaatctt atgtactctt atc 763

[0262]

<210> 149
<211> 518
<212> DNA
<213> Homo sapiens
<400> 149
gtcccgtac caqcgcacaga ctttgcacaa ac atgcagatgg ttctcacatq tttcccttgt 60
ctcattttca gggcacgtgt cctaggttct ttgcatttgc tcttcgaagg caaggtttcc 120
agatctctct gtatccatcat gttcccttt tggatgcacc ttaattttaa aatacctctt 180
tttctcatta attagatcac ttcaaggtaa atacaaaaca tggcaagatq gatttaaattt 240
taaqgggata taagtataca taagagaaga ccaatctta cttttaaaaa tgcagttaat 300
taacaataaa gtaaaatata gtgaaggatc tatacgctta tttgaacttt gctgttgtt 360
aggacttgcata ttccctctt aactttttt aaaaaactcg ttttqaaaac tagcatttgg 420
tgtgggtttc caatgtatca gaaacatgtt aactacaaca taatcggtt aagaattttaa 480
aataattttq caqtagaqaa tggtaatgtt ttccacccq 518

[0263]

<210> 150
<211> 334
<212> DNA
<213> Homo sapiens
<400> 150
cgcttaqctc cataqcccc taaccccatt tttaactqa gtcagaccct aactcaagaa 60
ggatgagttt aatgacttct aaagtttctt tctaatcctt tcacagtttq atttqactcc 120
ttcttaataaa gtcagtcgq actgataatq tgaccaccc aatactaccc taggagtcaa 180
tagcacacat qaqctaaaac ctqgtqgtct ctcqgctccc aqgtgtqgqac tqtcaccaqg 240
actatattct tcacccctttt cacaaccq aactqctqct qggcaagccca agaaaqcgaga 300
aaaaaaaaatcacttacttqgtttt tggccaaatqct qatc 334

[0264]

<210> 151
<211> 304
<212> DNA
<213> Homo sapiens
<400> 151
cgctgtataq ctaacaqcat aqccaatgcc ctqatcttct aqccacactq taatttcgtq 60
tccttgatc acaaccccaq ctgtatgtcat caaagtggaa aaactgacag cattgtctqa 120
aaqacacagc ttgggaaaat tgctccctgt aqctccccgt tcaatqaata ttatcatatt 180
tttctcatgg ataaqgttac tqtgtttttt ttqctgggcc qqqqgtqoooo agqcttctac 240
atacttagct actaaggcaca ttaccattcc gtttqtgaa tggactaaac qtgacaatgc 300
qatc 304

[0265]

<210> 152
<211> 304
<212> DNA
<213> *Homo sapiens*
<400> 152
cggcatacag gccataatTTT ccatccatga aaatgaaaaaa tggcgatgca agctatactt 60
tgtctacaat cacaacta^Q gtgaattctc aqgtggctct ttagagtctc tgaacagtt 120
atctttttag aatcaatca aaataatcca tggctcqagt tcagtttggc caqcacacca 180

131 agaqcagggtg gttaaaaaca ttcggactca atgctttca caatacataat tcagaggatg 240
aatctttaca caaatccac tacctctatt ctaacagttc tgacttggtt catccaatgc 300
gatc 304

[0266]

<210> 153
<211> 730
<212> DNA
<213> Homo sapiens
<220>
<221> variation
<222> 2, 10
<223> n is a or c or g or t.
<400> 153
tnacacaggn gatcgacatt gtqogggaggt gaggccaaqgc agctgggaca ttctagaat 60
qccttttag aattatggca ttctaaaaaa qgtacttcca gaacaqttgt tttagatggaa 120
ttcaaaagtaa agttcaqatt agcttttggag qaaaaaggaa tgctaccqta aatgtgaacc 180
ttaatccagt atactccqgc catccttqg qgatqaataa atttcaqgtt ttacctaagt 240
cttaacaatt tctttactag ccccttaact gttactttag caagtttgat aaaatgtgt 300
aaaaatataqgc atttttqagt ttttaacttt aqtgqgaaqgt agatqttca tcaacttctt 360
tggaaaataa agacactqgc atataaacta tattttttq agacaggatc tcactctgtt 420
gtccaaqctg qaatqcapta atgcaaaatc ggctcactgt agcctcgacc tcgtgggctc 480
aagcaatccct cccacccatcg cctccgtact aqtgaaacca caqacatqag ctqctqacc 540
cagctaaaaat qgagttttt taatttttgg gtctttaaa tgcatttqga ggtctttagt 600
tttacccac taaaaatttggg attttaattt aataaatca aqatgtgaa ctttacccac 660
attttaaaagc cattatattt ttctataaa ccctgttctc gtttggagga gaaaagaaatt 720
qgaattttcq 730

[0267]

<210> 154
<211> 509
<212> DNA
<213> Homo sapiens
<400> 154
caattgtacc atctagttt tttaaattat aattaattaa tttgtttttt gagaccaagt 60
ctcactctgt tgccccaggct ggagtcagg ggcatgatct caqctcactg caacctccgc 120
ctccctgggtt caagcaattc tqtgcctcag cctcccatgt aqctgggatt acaqqtgtga 180
qcaaccacgc ctggcttaatt tttgttatttt tagtagagac caggtttcac tatgttggcc 240
aggctggctc cqaactccctg acctcaggta atccgttcgc cttggccctcc caaagtqccg 300
ggattacagg cgtgagccac aqgtgccccgc caccatctag tttttaaaca qaaaccacaa 360
qcatctagcc cctgctggag qaaacttt aattttctt tccatatatq tqtqgtatqq 420
gtctttccctg qtgacatatt aagtgggttag tacttgggaa aatccaaac tattaagtt 480
qgggttttqa actctggacc cqacgtttq 509

[0268]

<210> 155
<211> 487
<212> DNA
<213> Homo sapiens
<400> 155
gatgcattg ccacaatcca gcctggatga caqagtqaga ccccatctcc aaaaaaaaaa 60
aaagaaatta qctqgcgcgt qatqgtgtgg ctqaqqqagg ctcaggtqgg agqattqctt 120
qaactcaggat qtctqaqact acagtoaqact agactctccct cactgcactc caoqcttqaa 180

133

gataagagcaa gaccctgtct cttaaaaaaac aggcaaatta ttgattataq ttcaagtggc 240
 taaatccctgt ggtagaaaca ttcatqcata ggggtggtqca tccacatcq tcttaggtq 300
 gaggaggqtc atggaaaggca gatttttgtg tgcacttgaa ctgaactttg aaggatgatt 360
 aqqaqtaqc tgggtggggg aaggagaaga cctaccaggc agaaggactg qcataagtaa 420
 aqgcttattt gtaagaagta qcatggcaca ttaagggaaag ttqaggttqtc tcaqtatgac 480
 cctaccq 487

[0269]

<210> 156

<211> 333

<212> DNA

<213> Homo sapiens

<400> 156

cgtcacacca acactgagaa ggtagttttt tcatttttc tgaggctcq agttaaata 60
 qctttcccgaa gatcacacat tggcatgctg ggtttqaac ccagagqtc tqaatccaga 120
 qctctqaactt cgtatccctg cctcaqccct tcaagttqct gggagagaat tttttaaattt 180
 ttgttagaaa caagggtctca ctacattgcc cacactgqtc ttqaactcct qgcctcaagc 240
 aatccctccca ctttqccctc ccaaaggqct gggattacag gqgtqagtc ctgtgcccq 300
 aacagagcct qaaactttaa cgacaatqcg atc 333

[0270]

<210> 157

<211> 763

<212> DNA

<213> Homo sapiens

<400> 157

gatcgcatttgc tcaacaggctt gtccttaqtt ggggggtqat ctgtgttggg aaaccctttt 60
 agaaagaatt actcttgagc tgagattttt gcccgttta gtgtgtatcc ccagaattac 120
 tagtqaatq gtttqtaaa taqggttaatt tttttccct cctgaatttq gtggatcttt 180
 tagagttgaa tagacacqac tttctgcaac agatqggqat ttagacacqg aggacacttt 240
 agttgtqctt agtttcttct accatccgtq qcatggaaat aatttaggtat tataatgatt 300
 taaaggactt aagttcagtc tttttgggtt ttttcttttccataaqtatttct tattaggtac 360
 gataaaagcat tgattgcaag tacttcttta accatacatq atqatgtatt ttcgtqaaag 420
 tagagctqta aaaatctqac ccattcagct gggctcaqgg qctcatqccct qtaatcccaq 480
 cactttggaa gqctqagqca gaaaggatcac ttgagatcq qagttcaaga ccagcctaag 540
 caacatqca agaccccatc tctacqaaa atttttttttaa ttagctggac atqcccctgt 600
 agtcccactc aggaqactq ggttggqaaqca tcacttgcac ccqggaggc qaggctacag 660
 tgagctgtqat ttgtqccact qcaactccac ctgagtgaca qagcaqacc ctgtctcaaa 720
 aataaaaataa atctqactca ttaaattqat ttttaataac tcg 763

[0271]

<210> 158

<211> 470

<212> DNA

<213> Homo sapiens

<400> 158

tttgattqcc cattacaaca acttggagggt tagaaaatata atattgatata taagtccctc 60
 cccagacaga ttatattaa gtcctccqta gggqctttct ccccaqgggaa cattattttq 120
 tagccaggc tgagaatcac tgcattaaagg acaaattggac tagggqatataa aaagctcaaa 180
 aqttggqcttq caqatataq gtttqatqatc acatqctqtt tacctctaaa tttatataat 240
 aattttgaaat atttggaaaaaa gataattact tagttcaaaa tttaatttag ataaatttcc 300
 gctataqctt aaaatqataa ctgtctatg tctcagaata attaattttq gctqggqctq 360
 gttggctact cctgtaaaatc ccaqccccc gggaggccaa ggttgggtqga tcacttqagg 420

131

132

agagcaggta gttaaaaaca ttcggacta atgtttca caatacaaat tcaaggatq 240
 aatcttaca caaatccac tacctctatt ctaacagtgc tgacttggtt catccatgc 300
 gatc 304

[0 2 6 6]

<210> 153
 <211> 730
 <212> DNA
 <213> Homo sapiens
 <220>
 <221> variation
 <222> 2, 10
 <223> n is a or c or g or t.
 <400> 153

tnacacaggta gatcgactt gtggggagggt gaggccaggc agctggaca ttctagaaat 60
 gcttttttag aattatggca tttctaaaaa ggtacttcca gaacagtgtt ttagatggaa 120
 ttcaagtaa agttcaqatt agctttqag qaaaaaggaa tqctaccqta aatgtgacc 180
 ttaatccagt atactccagc catccttgg ggtatgaaata atttcaqgtt ttacctaagt 240
 cttacaatt tctttactag ccccttaact gttactttag caggttqat aaaatgtgct 300
 aaaaatatgc attttqagt ttttaactt agtggaaagt agatgttca tcaacttctt 360
 tqggaaataa agacactggc atataaactt tatttttg agacaggatc tcactctgtt 420
 gtccaagctg gaatgcaqta atgcaaacat ggctcactgt agcctcgacc tcgtgggctc 480
 aagcaatcct cccacccatq cctccctqact agtggaaacca cagacatqag ctgctgcacc 540
 caqctaaaat qqaqtatttt taatttctgg gtctttaaa tgcatttggg qgtctttqgt 600
 ttacactc ac tggaaatttagg attttaattt taaataatca aagatgtgaa cttacagac 660
 attttaaagc cattatattt tttctataaa ccctgttctc gtttggagga gaaagaaattt 720
 ggaattttcg 730

[0 2 6 7]

<210> 154
 <211> 509
 <212> DNA
 <213> Homo sapiens
 <400> 154

caattgttacc atcttagttt tttaaattt aattaattaa tttttttttt qagaccaagt 60
 ctcactctgt tgcccaagggt ggagtccagt ggcatgatct cagtcactg caacctccgc 120
 ctccctgggtt caaqcaattt tqtqccctcaq cctcccatgt agctgggatt acaqgtgtqa 180
 gcaaccacgc ctggcttaatt tttgtatttt tagtagagac caggttcac tatgttggcc 240
 aggctggctc cgaactccgt acctcaqgtt atccgttcgc cttggctcc caaagtgcgg 300
 ggattacagg cgtqagccac agtggccggc caccatctgt tttttaaaca gaaaccacaa 360
 gcatctagcc cctgctggag qagaaacttt aattttctt tccatataatq tqtgtgtgg 420
 gtctttccgt qtgacatatt aqgtgggttqg tacttggggaa aatccaaaac tattaqttt 480
 ggggttttga actctggacc cgacqtttg 509

[0 2 6 8]

<210> 155
 <211> 487
 <212> DNA
 <213> Homo sapiens
 <400> 155

gatcgcatgt ccacaatcca gcctggatqa caqagtqaga ccccatctcc agaaaaaaaaa 60
 aaagaaattha gctggccgt gatgggtgtgg ctqagggagg ctcagggtgg aqgattgctt 120
 gaggctcaqgt qtcgtqagact acagtqagct aqgctctcct cactgactc caqctggaa 180

133 134
 gataqaqcaa qaccctgtct cttaaaaaac aggcaaatta ttgattataq ttcaqgtggc 240
 taaatcctgt ggtagaaaca ttcatgcata ggggtggtgc tccacatcaq tcttaagggtgg 300
 gaggagggtc atqqaaggca gatttttqgt tqcacttcaa ctgaactttq aqgatgatt 360
 aggagtaagc tgggtggggg aqggaagaa cctaccaggc agaaggactg qataagtaa 420
 aggcttattt qtaagaagta qcatggcaca ttaaggaaq ttgaggttgc tcagtatgac 480
 cctaccg 487

[0 2 6 9]

<210> 156
 <211> 333
 <212> DNA
 <213> Homo sapiens
 <400> 156
 cgtcacacca acactqagaa ggtqatattt tcatttttc tgaggctcaq agtttaata 60
 gcttcccgaa gatcacacat tggcatgctg ggatttqaac ccqagagtc tqaatccaga 120
 qctqaaactt cgatcctctg cctcagccctc tcaqtaqct gggagagaat tttaattt 180
 ttgttagaaa cagggtctca ctacattgcc cacactggc ttgaactcct ggcctcaagc 240
 aatcctccca ccttggccctc ccaaaggqct gggattacag gggtagtca ctgtgcccag 300
 aacagagcct qaacttttaa cgacaatgcg atc 333

[0 2 7 0]

<210> 157
 <211> 763
 <212> DNA
 <213> Homo sapiens
 <400> 157
 gatcqcattt tcacaggctt gtccttagtt ggagggtgat ctgtgttggg aaacctttt 60
 agaagaattt actcttqagc tqaqattttt qccctqttta qttgtataatc ccagaattac 120
 tagtqaaatq ggttqtaaa tagqgttaatt tttttccctt cctqatttq gttgtatctt 180
 tagatqgaa tagacagcac tttctqcaac aqatqgggat ttagacaggg aggacacttt 240
 agttgtgctt agtttcttctt accatccgtg qcatqaaat aatttaggtat tataatgatt 300
 taaaggactt aqtttcagtc ttttggttc ttttcttataaagattt tattaggtac 360
 qataaaqcat tqaattqcaag tacttcttta accatacatq atqatqattt ttcqtaaq 420
 tagatqgta aaaatctac ccattcaqct gggctcaagg gctcatqct qtaatcccaq 480
 cactttqgaa gqctqaggca qaaggatcac ttgagatcag gagttcaaga ccagcctaag 540
 caacatagca aqacccttc tctacagaaa attaaaaaa tttagctggac atqcccctgt 600
 aqcccactc aqgagactgq ggtgggagga tcacttqagc ccaggaggc qaggctacag 660
 tgagctgtqa ttgtgcccact qactcccaqct tqagtgaca qaqcaqacc ctgtctcaaa 720
 aataaaataa atctqactca tttaatttqat tttaataatc tcg 763

[0 2 7 1]

<210> 158
 <211> 470
 <212> DNA
 <213> Homo sapiens
 <400> 158
 cttgattgcc cattacaaca acttggggat tagaaaaat atattqat taaqtcctc 60
 cccagacaga ttatattqa qtcctccaga ggagctttctt ccccaaggaa cattattttq 120
 tagccagggtc tgagaatcac tqcattaaqg acaaattggac taaaqgttaaa aqgctcaaa 180
 agtggqcttg cagatttqa gttgtqatc acatqctgtt tacctctaaa ttatatttaat 240
 aatttqaaat atqggaaaaa qataattact tagttcaaaa tttaatttq ataaatttcc 300
 qctataqctt aaaatqataa ctqatctatq tctcagaata attaattttq gctgggggtg 360
 gttggctact cctgttaatc ccqgacccct qggaggccaa ggtgggtgaa tcacttgagg 420

135

136

tcaggagttc gagaccagcc tgcccaacat ggtgaaaccc cgtctctacg

470

[0272]

<210> 159

<211> 295

<212> DNA

<213> Homo sapiens

<400> 159

tggattggtc ggggtgtggc gctcatgcct gtaatcccaq cactttggga gccccaaqgtg 60
gggtggatcac ctaagggtcag gagttcaaga ccaggctgcac caacatggag aaaccccatc 120
tctactaaaa aaaaaaaaaa aaaaaaaaaaqt tggcatgtt ggtgcattcc tggatccca 180
gttcctcggg aggtctggcc agqaaatcgtt cttaaaccq gqagccaaqg attgcagtga 240
ccaagatcac atcattgcac tccagcctgg gcaacaagag cgaaactcca tctcg 295

[0273]

<210> 160

<211> 295

<212> DNA

<213> Homo sapiens

<400> 160

gttgcgtatcc cagatactcg ggagggtgaaq gccccaaq cacttgcacc tggggaaagg 60
cggttgcgtt gqccqatcat cccaccq cactctgcctt tggggcaacq tggactcgg 120
tctc 124

[0274]

<210> 161

<211> 608

<212> DNA

<213> Homo sapiens

<400> 161

caaacgtcgq ggcattcccgq ataggcccgq aaaaqtgttq gggaaqaaq ttagatttac 60
ggccgatgaat atgatagtqaa aatggatttt ggcgttagtt tggctcaqgg tggatgcctga 120
gaaataqggaa aatcaqgtqaa tggatgcctcc tatqatqgca aatacaqctc ctattqataq 180
gacatagtgg aqgtqgqcta caacgtatqta cgtgtcggtq aqgtqatgtt ctatgtatq 240
gtttqctaat aacaatqcccg tcaqgcccacc tacqgtgaaa agaaagatqaa atccctaggc 300
tcagagcact qcagcagatc atttcatatt gcttccgtgg agtgtggcga gtcagctaaa 360
tactttgcg ccgggtgggaa tagcgtatqat tatggtagcgt gagggtqaaat atgctcggt 420
gtctacqctt attcctactq taaatatatq tggatgcctac acgataaaacc ctagqaagcc 480
aattgtatc atagctcaga ccataccatq gttatccaaat ggttctttt ttccggagta 540
gttaqgttaca atatqggqaa ttattccgaa gcttggtaqgq ataaqaaat ataaatcagg 600
gtgaccqg 608

[0275]

<210> 162

<211> 315

<212> DNA

<213> Homo sapiens

<400> 162

gatctcagac cacagtggaa taaaactqga aatcaactcc aaaaggaaacc ttcaaaacca 60
tggaaataaa tggaaattaa ataacctccctt qaattatcat tgggtcaaca gaaaaatcaa 120
gatqgaaattt taaaattqaa ctqaacaata ataqgtacac aaccttatcaa aacctctqgg 180
atacaqcaaa gqcaatqcta agggqaaagg tcacagccta aatgcctaca tcaaaaagta 240
tggaaqaaca caaataqaca atctaaqgtc acacctcaqg qaactqaqaa aacaqaaca 300
aaccaaaccg agccc 315

[0 2 7 6]

<210> 163
 <211> 666
 <212> DNA
 <213> Homo sapiens
 <400> 163
 ctatccatg ctataaagt gacgaaccctt gcaaacccta cgtgaaatga aaaaagccaa 60
 gacccaagag gtcacctata gtagtggatca acttggatga aatatccaga atagaqaaat 120
 ccatagagac agaaaggaga cactggatg tgccagggt tggaqagagg aaagaatggg 180
 aagaqagact acataqatqag taagggttt tactttggag tgatgaaat gcttggatc 240
 tagatagaqg tgccgggttt gacacaacgtt gtagatggac taaatgccc taaattgttc 300
 actttatgtt tttatttat ttatgtttag agacggggtc ttgctgatt 360
 gcccaggctg ctgtcaaact cctaaqctca cccaggccca atggctcatg cttgtatcc 420
 cagcactggg aqgccaagg gggcgatca cctggatca qgagttcgag accagcatgg 480
 ccaacatggc qaagccccgt ctctactaa attacaacaa ttatgtttag tggtggatq 540
 acggctgttag tcccaqctac ttagggatct gaggcaggag aatcgcttga accccggagg 600
 ccqaaqgttq cagtqagccq agatcactt aqctggccq acaagagcaa 660
 gactcg 666

[0 2 7 7]

<210> 164
 <211> 201
 <212> DNA
 <213> Homo sapiens
 <400> 164
 caaacgtcgq gtcacatggg aatgtqgtaa aqccctatta ctgtqaggtt cagacatgt 60
 cactaaaaaa qagctttatt gggccgggqcg cgggtggctta ctccataat cccagcactt 120
 tggggggccq aqgcaggcaq atcactgtqg qtcaggatg cgggaccacg ctggctaaaca 180
 tggcaaaacc ctatctttc c 201

[0 2 7 8]

<210> 165
 <211> 340
 <212> DNA
 <213> Homo sapiens
 <400> 165
 gatcaatcgc ccacatcaq cqaqacgtaa attatqgctg aatcatccgc tacccacq 60
 ccaatggcgc ctcaatattc tttatctgcc cttccatca catcgacqg ggcctatatt 120
 acggatcatt tctctactca gaaacctgaa acatcgatc tattccctg cttgcaacta 180
 tagcaacacg cttcatggc tatgtccctc cgtqaggcca aatatcattc tgaggggcca 240
 cagtaattac aaacttacta tccggccatcc catacattg gacagaccta gttcaatgaa 300
 tctqaggqgq ctactcaqta gacagtccca ccctcacacq 340

[0 2 7 9]

<210> 166
 <211> 578
 <212> DNA
 <213> Homo sapiens
 <400> 166
 gatcgcattq ctataaagaa ctatctqaga ctggtaatt tatqaagaaa aqgqgttta 60
 attgacttac agttcccttag gctgtacaq aqccctggct ggggagattt caggaaattt 120
 aaaaccatgg cggaaaggcga aqgggaaqca ggcattgttt atatgacctg aqaaaggagga 180
 gqagqgtqca gggggggqgtq ctacacactt ttaaacaacc aqatctcgqg agaactcact 240

139

accacaagaa cagcaaggaa caactctgtc cccatgatcc aqtcacctct caccagaccc 300
 ctccaaact ggggattaca attttacatq agatttgggt gaagacacaa atccaaacca 360
 tgcgccccatq gttaaagccac acatgactgt atataaccta cttgtatcct tgcataact 420
 ttaaataatc tcttagattac acattataat acctaataca ggcgggtgt ggtggcccat 480
 gcttgtatc ccaqacttt caqaggccaa ggcagggtgaa tcacqaggc aqgqgttcaa 540
 gaccagcctg gccaacacag tgaaacccca tctctacg 578

[0 2 8 0]

<210> 167
 <211> 143
 <212> DNA
 <213> Homo sapiens
 <400> 167
 gtttcgcag tcctgaacaa aatactaaca aactgaaactc agcaatataat gaaatggatt 60
 atatgccccatq aaaagttaga tttatataa gaaacggcaag gtcgggtctt catacaaaaa 120
 tcaatttatt acaccatatt aac 143

[0 2 8 1]

<210> 168
 <211> 637
 <212> DNA
 <213> Homo sapiens
 <220>
 <221> variation
 <222> 39
 <223> n is a or c or g or t.
 <400> 168
 gatcgcattt gcaagttaaa tccttaaqua aaaaagganc aaagctggag gcatcatgct 60
 acctgacttc aaaactatac taccaggctt ccaqtaacca aaacagcatq atactggatc 120
 caaaacagat atataqacca atggaaacaga acagggcct caqaaaataac accacacatc 180
 tacaatcatc tgcattttga caaacctgac aaaaacaaga aatqgggaaa gaattcccta 240
 tttataataat ggtgttggga aaactggctt gccatatgta gaaaqctgaa attggatccc 300
 ttcccttatac cttataaaaa aattaattca agatggatc aagacttaaa tggtagaccc 360
 aaaaaccataa aaaccctaga agaaaaccta ggcqataccct ttcaggacat aggcatgggc 420
 aqgacttca tgcataaaac accaaaagtg atggtaacaa aqgcaaaaag ccaaaataga 480
 caaatqggat ctaattaaac taaagqctt ccgcacacqca aaqaaaacta tcatcaqat 540
 gqgcaqacaa cctaaqagt ggggagaaaat tttqcaatc tacacatttq acaaaggqct 600
 aatttccaga atctacaaag aacttaaaca aatttac 637

[0 2 8 2]

<210> 169
 <211> 433
 <212> DNA
 <213> Homo sapiens
 <400> 169
 gatcaatcgc atcatcaccc aqctgtcaa aaatgtqaac attcgtggat tttttctat 60
 ttgcaqcca aaaaatqaaq actgtcaact tcatggatata taaaagaagg catqagattc 120
 ctggcctga taaaaatata ttgatggttc atcacacaaa tggaaqcaag agctccatata 180
 ctgtatctgt tccccacqtc cactccaqgt qacaqggatq atcaaqaqaa cccaaactqqa 240
 tqctttctgc tcaqtagatt ttgatcacaq ctgaaqqaatc tcaaaqgtttaa qaaatcccaq 300
 tqctttqgca gggcacqgtq gtcacqccct qtaatcccaq cactttqqa ggcggqgcq 360
 gqgagatcac qagggtcaqqa gattgagacc atccctggctt acacqgtqaa accccgtctc 420
 tactaaaaat acg 433

【0283】

<210> 170
 <211> 739
 <212> DNA
 <213> Homo sapiens
 <400> 170
 ctgatccatg accggtaaa aatgtcttq aaatcacatc aqgctcatca aqgacattta 60
 tatgtgaga gagacctaga gaaagatgat gacctgaaac ctgaaaggaa gtctttatg 120
 tagagactct gcctatgaca gtgaccccgta gtagactact gttcacagga gcaagtcct 180
 ggaaccaacc caaatgccc tcaacgataq actggataca gaaatgtgaa atgcaccata 240
 tgcctcata tqaaggatctc tgcaaggatgg tctgactcac gcgagtggtt ttgagataaq 300
 ttgtgttata cctcaggta ggttaaggctg aqagatgaa taactaaaa aqtagtqatg 360
 ttacataga atctaaattt gaaaagacaa taatcagc attaaaaagt ttgcctttaa 420
 gtaaaaaaaa ttatqaqcca taataggtt tacattttc gtaatagaat atcctaaagc 480
 aataaaaccc ccaactatgtg aactctaagt tctqaatatac cgaattttc aaatagttct 540
 gagtttattt taatggttc ttactgact aqaaacacttc ctaattttc acatcatttt 600
 taacctcctt caaacataact gaatataattt tttccagggtt ttatttgagg tataattgac 660
 aaacaaaaat tqtgtatgta aggtgtacaa tqtgtatgaaa tqtgtatgtgaa aatacatttq 720
 atggattctt gggcttcctt 739

【0284】

20

<210> 171
 <211> 380
 <212> DNA
 <213> Homo sapiens
 <400> 171
 gatctgacac caagcacagg caacagaagt gaaaataqgc aatggact atatcagatg 60
 taaaatgttt tgcccttaaa taacacaaca attqacaaq tcaaaagaca acctacqgaa 120
 taagagaaaa tatttgcaaa tgatatactt gatatgtgtt taatattgaa aatatacaca 180
 gaactacaac aacttaacac caaaaaatc aaataacaca attaaaaat gggcaaaagag 240
 cttgaataga cacgtctcca aagaagatata acaaattgaca aacaagcactg tgaaaaatg 300
 ttcatttata atcattaaqaq aaaqcaatt tqaaccgca atqagataacc aactcataacc 360
 tatacgatg gctactatcg 380

【図面の簡単な説明】

【図1】この発明の実施の形態1である遺伝子発現プロファイル解析装置の構成を示すブロック図である。

【図2】遺伝的アルゴリズムにおける遺伝子の組み合せの表現方法を説明した説明図である。

【図3】遺伝的アルゴリズムにおける遺伝子の組み合せの評価関数（予後の存命月数の予測誤差に関する評価）のグラフの一例を示す説明図である。

【図4】遺伝的アルゴリズムにおける遺伝子の組み合せの評価関数（採用した遺伝子数に関する評価）の説明図である。

【図5】従来の方法によって予後の存命月数を推定した結果を示す説明図である。

【図6】この発明の方法によって予後の存命月数を推定した結果を示す説明図である。

【図7】各遺伝子と予後の存命月数の相関係数の分布を示した説明図である。

【図8】この発明の実施の形態1である遺伝子発現プロファイル解析方法の処理の流れを示すフローチャートである。

【符号の説明】

1	入力手段
2	データ前処理手段
3	遺伝子組み合わせ手段
4	第1推定手段構築手段
5	第1推定手段
6	遺伝子組み合わせ評価手段
7	組み合わせ制約手段
8	第2推定手段構築手段
9	第2推定手段
10	出力手段
S1	入力ステップ
S2	データ前処理ステップ
50 S3	遺伝子組み合わせ候補生成ステップ

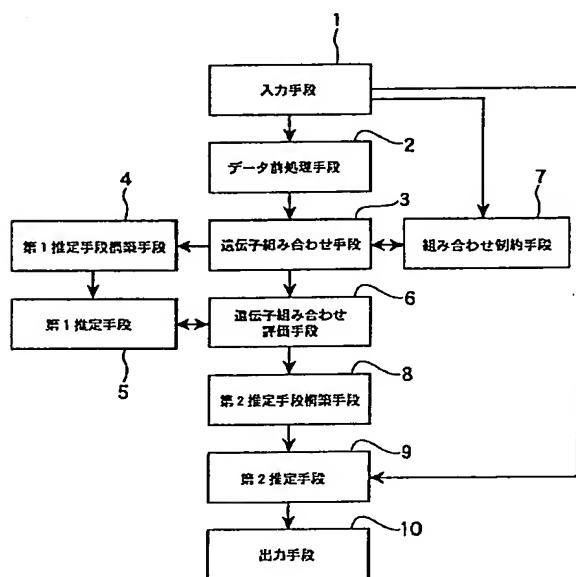
143

S4 第1推定方法生成ステップ
 S5 第1推定ステップ
 S6 遺伝子組み合わせ評価ステップ
 S7 遺伝子組み合わせ抽出ステップ

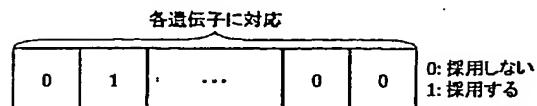
144

* S8 遺伝子組み合わせ制約ステップ
 S9 第2推定方法生成ステップ
 S10 第2推定ステップ
 * S11 出力ステップ

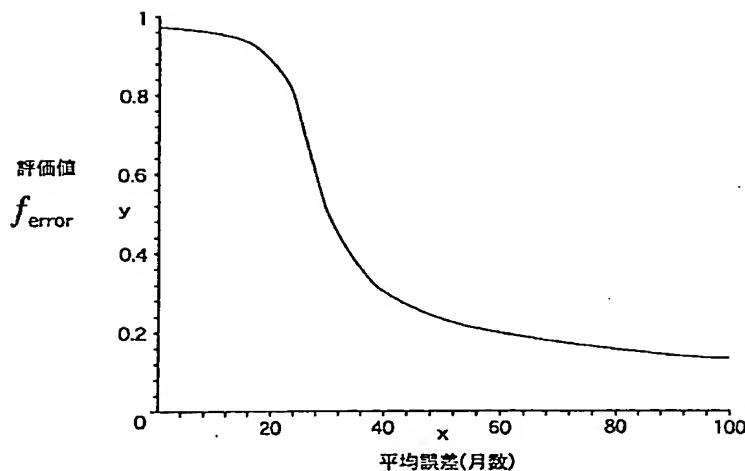
[図1]



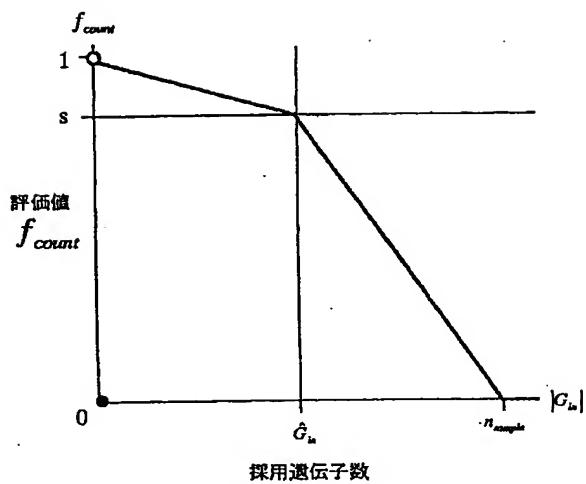
[図2]



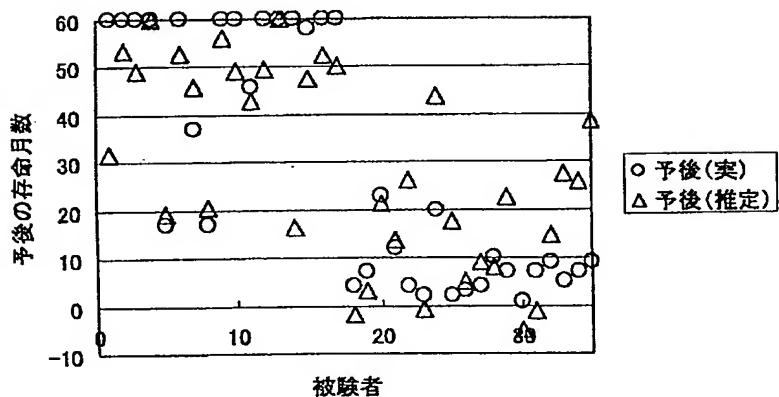
[図3]



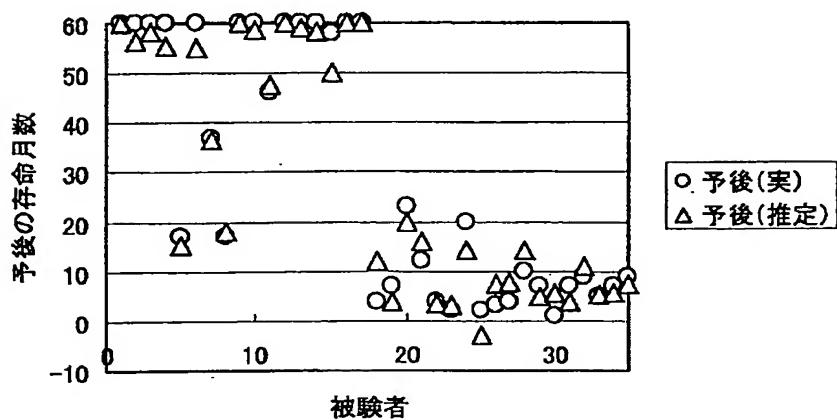
[図4]



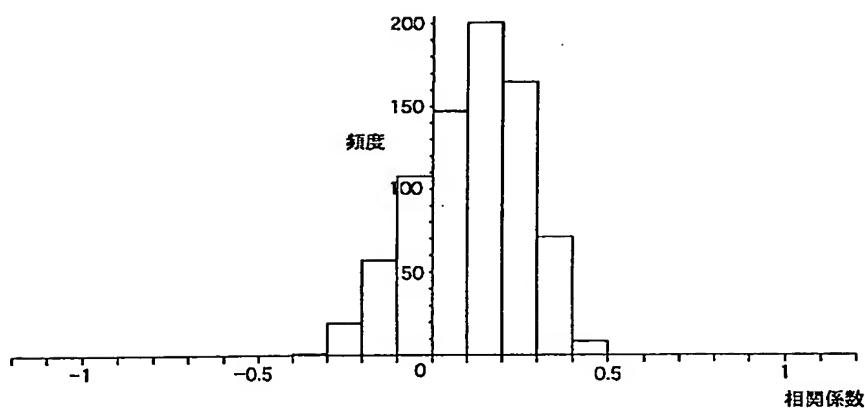
[図5]



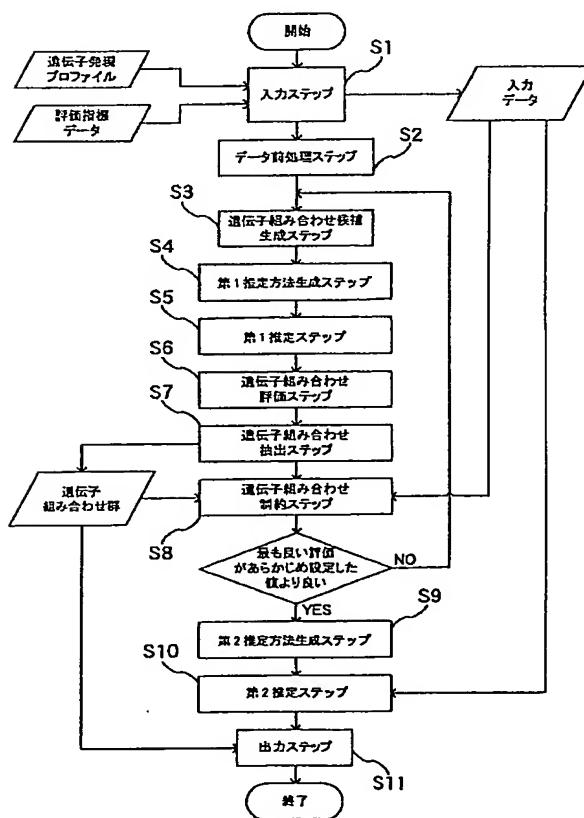
[図6]



【図7】



【図8】



フロントページの続き

(51) Int.C1.
 G 0 6 F 17/30
 G 0 6 N 3/12

識別記号
 1 7 0

F I
 C 1 2 N 15/00

マーク (参考)
 Z N A A
 F

(72)発明者 蜂野 純一
滋賀県大津市瀬田三丁目4番1号 寶酒造
株式会社中央研究所内

(72)発明者 向井 博之
滋賀県大津市瀬田三丁目4番1号 寶酒造
株式会社中央研究所内

(72)発明者 浅田 起代蔵
滋賀県大津市瀬田三丁目4番1号 寶酒造
株式会社中央研究所内

(72)発明者 加藤 郁之進
滋賀県大津市瀬田三丁目4番1号 寝酒造
株式会社中央研究所内

(72)発明者 香川 芳範
兵庫県尼崎市塚口本町五丁目4番36号 三
菱スペース・ソフトウェア株式会社関西事
業部内

(72)発明者 谷嶋 成樹
兵庫県尼崎市塚口本町五丁目4番36号 三
菱スペース・ソフトウェア株式会社関西事
業部内

(72)発明者 永野 隆文
東京都千代田区丸の内二丁目2番3号 三
菱電機株式会社内

(72)発明者 山田 訓
東京都千代田区丸の内二丁目2番3号 三
菱電機株式会社内

F ターム(参考) 4B024 AA11 AA19 AA20 CA04 CA09
CA12 CA20 HA14
4B029 AA23 BB20 CC03 FA15
4B063 QA01 QA05 QA19 QQ02 QQ53
QR32 QR56 QR62 QR82 QS16
QS25 QS34 QS39
5B075 ND20 NK02 PR06 QM08 UU19